TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

TOMMI HYNNINEN

PRODUCT DATA MIGRATION BEFORE INFORMATION SYSTEMS RENEWAL

Master of Science Thesis

Prof. Samuli Pekkola has been appointed as the examiner at the Council Meeting of the Faculty of Business and Technology Management on September 5th, 2012.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Information and Knowledge Management

HYNNINEN, TOMMI: Product data migration before information systems renewal
Master of Science Thesis, 101 pages, 17 appendices (29 pages)
April 2013
Major: Business Information Management
Examiner(s): Professor Samuli Pekkola
Keywords: Product data, data migration, information systems renewal, product data management, PDM, master data management, MDM, knowledge management.

Essential for organization's existence is to make profit by selling its goods and services. This is quite straight forwarded process, though it requires enabling and supporting functions in order to succeed. With these functions are referred to the organization's business processes and information systems, which in combined enable and support organization's sales. Organization's offering is constantly changing and respectively the amount of information related is increasing. In addition to the changes, the business requirements and ways of working may change as well. This leads to development of business processes and information systems, which has to adapt for the change.

Researched company faces challenges with product data, its management and information system renewals, and at the same time, compliance issues with its parent company. Thus, the research objective is to introduce a holistic model for product data management, which includes the known and upcoming information system renewals and as well considers the mentioned compliance issues. To create a holistic model for product data management within given requirements, a two-part research was conducted consisting of theoretical section and the empirical section. Theoretical section is based on scientific literature about research areas and the empirical section is based on qualitative data analysis, in which the analyzed data is received from total of 17 interviews. With theoretical and empirical sections, the holistic model is created.

Results showed that the product data is scattered, deficiency of information, lack of documentation and large quantity of tacit knowledge. For information systems, Product Register is obsolete, ERP-system will be replaced and Sinfos has altered requirements. Recommendations include that product data related functions to be revised, documented and trained properly. In addition, segregation of duties should be revised and person in charge should be named. Product Register should be replaced with a system consisting of product master data and linkage to the new ERP-system, and it covers the changed requirements for Sinfos. Upcoming ERP renewal has to be considered when dealt with other renewals and in information systems architecture. For further research, the effect and execution of implemented solutions for depicted challenges should be studied, in contrast to the theoretical literature.

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO
Tietojohtamisen koulutusohjelma

HYNNINEN, TOMMI: Tuotetiedon migraatio ennen tietojärjestelmien uudistamista
Diplomityö, 101 sivua, 17 liitettä (29 sivua)
Huhtikuu 2013
Pääaine: Tiedonhallinta
Tarkastaja(t): professori Samuli Pekkola
Avainsanat: Tuotetieto, data migraatio, tietojärjestelmien uusiminen, tuotetiedon hallinta, PDM, master datan hallinta, MDM, tietojohtaminen.

Yrityksen tehtävä on tehdä voittoa myymällä tuotteitaan tai palveluitaan. Myyntiprosessi on melko suoraviivaista, mutta se tarvitsee tukevia toimintoja onnistuakseen. Tukevilla toiminnoilla tarkoitetaan liiketoimintaprosesseja sekä tietojärjestelmiä, jotka yhdistettynä mahdollistavat yrityksen liiketoiminnan ja myynnin. Yrityksen tarjoama muuttuu alati ja vastaavasti tarjoamaan liittyvä tiedon määrä kasvaa. Tarjoaman muutosten lisäksi liiketoiminnan vaatimukset ja työskentelytavat saattavat muuttua. Tämä johtaa liiketoimintaprosessien ja tietojärjestelmien kehitykseen, joiden täytyy kyetä muuttumaan muuttuviin vaatimuksiin.

Tutkimuksen kohdeyritys on kohdannut haasteita tuotetietojen ja niiden hallinnan, tietojärjestelmien kehityksen sekä konsernin standardeihin mukautumisen kanssa. Tämän nojalla tutkimuksen tavoite on esittää kokonaisvaltainen tuotetietojen hallintamalli, joka sisällyttää tulevat tietojärjestelmäuudistukset sekä huomioi konsernin standardien rajoitteet. Kokonaisvaltaisen tuotetiedon hallintamallia varten, tutkimus toteutettiin kaksiosaisena, koostuen teoreettisesta ja empiirisestä osiosta. Teoreettinen osuus pohjautuu tieteelliseen kirjallisuuskatsaukseen ja empiirinen osuus perustuu kvalitatiiviseen tiedon analysointiin, jossa analysoitu tieto on kerätty 17 haastattelun pohjalta. Teoreettisen ja empiirisen osioiden perusteella kokonaisvaltainen malli luotiin.

Tulosten perusteella tuotetiedot ovat hajaantuneet, tieto ja dokumentointi ovat puutteellisia, ja hiljaisen tiedon osuus on huomattava. Tietojärjestelmistä Tuoterekisteri ei kykene täyttämään tarkoitustaan, ERP-järjestelmä tullaan korvaamaan sekä Sinfoksen vaatimukset ovat muuttuneet. Suosituksena tuotetietoihin liittyviä toimintoja tulisi tarkistaa, kehittää, dokumentoida ja kouluttaa työntekijöillä kunnolla. Lisäksi työnjako tulisi tarkistaa ja nimetä vastuuhenkilö. Tuoterekisteri tulisi korvata tietojärjestelmällä, joka sisältää yrityksen tuote master datan ja liitynnän uuteen ERP-järjestelmään sekä kykenee vastaamaan muuttuneisiin Sinfoksen vaatimuksiin. Tuleva ERP-järjestelmän muutos ja sen vaikutukset tulisi huomioida muissa uudistuksissa sekä tietojärjestelmäarkkitehtuurissa. Jatkotutkimusta varten käyttöönotettavien ratkaisujen vaikutusta ja toteutusta tulisi tutkia ja verrata löydöksiä teoreettiseen kirjallisuuteen.

PREFACE

“*Opportunity is missed by most people because it is dressed in overalls and looks like work.*”

- Thomas Edison -

As the studies at the university may be too theoretical every now and then, the practical work in real companies and especially with this master’s thesis, can and will be an eye opening experience. There is a considerable contrast between theoretical know-how and the practical learning by doing. Thesis process has taught me a lot and it was elevating experience as one could apply the gained theoretical knowledge into practice. Regardless of the valuable experience and adapted knowledge, I have to admit, that the writing process was not so delightful, though I can gladly state that: *“here we are with complete master’s thesis”*.

I would like to state my acknowledgement for my supervising Professor Samuli Pekkola, who gave feedback, advice and improvement ideas for my thesis and challenged me to consider my choices and ideas, instead of dictating what I should do. This has expanded my way of thinking and my skill of judgment for comparison, and for this I am grateful.

For my supervisor and project sponsor, Andrew Pickersgill, from the case company, I would like to express my gratitude and appreciation. First, you made this possible by providing the opportunity for thesis and second, you taught me a lot about practicalities and regularities of business world, and provided an insight to the functioning of globally operating organization. Also, I would like to dedicate my acknowledgement for my co-workers for great and supportive work environment, and especially for the interviewees, who made this research possible by providing your time and effort for the interviews.

Lastly, I would like to thank my family for the support and encouragement along with my thesis process, and a word of gratitude for my friends, who reminded about the life outside the thesis and data migration.

Turku, April 4th 2013

Tommi Hynninen

TABLE OF CONTENTS

ABBREVIATIONS AND NOTATION

ERP	Enterprise Resource Planning
IS	Information System
IT	Information Technology
ITF	Imperial Tobacco Finland Oy, the case company of this research
ITG	Imperial Tobacco Group PLC, the parent company of Imperial Tobacco Finland Oy
MDM	Master Data Management
NAV	Microsoft Dynamics Navision, the current ERP-system of Imperial Tobacco Finland Oy
PDM	Product Data Management
PLM	Product Lifecycle Management
SAP	SAP, the upcoming ERP-system that will replace the Microsoft Dynamics Navision

1. INTRODUCTION

1.1. Background

The core reason for organization's existence is to make profit, excluding the non-profit organization obviously. The function of making profit is dealt with the offering of organization's goods and services. In this research, case company Imperial Tobacco Finland Oy, is focused on selling its selection of goods. Though the selling of the goods is quite simple, the processes and required functions behind the selling are the issues that make the whole thing interesting and a bit complicated. Each product has certain data related to it, and the product data content may vary drastically from product to another. Therefore the management and handling of product data is crucially important.

Kropsu-Vehkapera et al. (2009, p. 759) describe the management and handling of product data as product data management, which objective is to help companies to manage their product related data and functions, and operations electronically, resulting with better efficiency and functionality. They emphasize, that without proper product data management, the control of the products will be lost which will result as lost sales and bad reputation. For different scopes of data, the data can be seen as two major categories of master data and transaction data. Master data is the data that changes on rare occasions and is mostly static, whereas transaction data is related to rapidly changing data (Loshin 2009).

With the vast amounts and constant information flows of product data, there has to be a tool for its management. Nonaka & Takeuchi (1995) present the concept of knowledge management which can be implied to organizations functioning by focusing on knowledge creation and sharing. From the technical perspective, information systems bring relief for ever-growing burden of data storing and processing.

Though, the knowledge management practices and supporting information systems would be function along, there may rise occasions where the new business requirements demand changes in organizations functions. New product data is needed and information systems will become obsolete. Then, the organization will be facing data migration where the data from old systems is transferred to the new systems. Also this requires product data quality assurance and refinement. (Thalheim & Wang 2013) The situation within Imperial Tobacco Finland Oy is as described above and it is the studied in this research.

1.2. Research problem and research questions

The case company, Imperial Tobacco Finland Oy (*later ITF*), is going through compliance processes and tries to fulfill given compliance standards and regulations by its parent company, Imperial Tobacco Group PLC (*later ITG*). The main problem with the compliance regulations is the historical burden of ITF. Before merging with the ITG, ITF had gone through several mergers and as a result of several mergers, the IT infrastructure and the information systems were established and enabled by different solutions, some as ad-hoc solutions. Neither with this tangled network of different systems and platforms it is not easy to neither maintain, update nor change those solutions. As a real life example, before merger with ITG, the ITF got through data migration with its previous parent company and after that began the new data migration with ITG. Quite complicated as it sounds, the ITF has managed its way with existing problems.

The next big thing is to renew ITF's enterprise resource planning system (*later ERP system*) and take a step towards fulfillment of the compliance regulations. Though the ERP system is an application or software among others, it is quite extensive and most comprehensive system within ITF. With the upcoming ERP renewal, there is a simultaneous opportunity to adjust and renew the IT and IS architecture as well. The reason for other renewals is the historical burden of old and out of date applications, and the new requirement changes within Sinfos product data bank application, which provides product data for wholesalers.

As the information systems and the architecture is tangled so is the product data and its management. Within several applications there exist duplicate or multifold data representations of the same data in different systems. Due to scarce and lacking documentation and the reality of limited resources as in manpower and time, this research delivers a solution for described problem. Also, to make things complicated, the existence of ITG, the parent company, sets some limitations and standards to be followed by ITF. By the given research problem, the main research question for this thesis is: *'How company should arrange its product data management?'* To answer this main question, there are four supporting research questions derived from the main question. The supporting questions are:

1. In which information systems the current product data is located?

2. In which form the current product data is available in the information systems?

3. What kind of product data is needed for renewed information systems?

4. How product data management should be arranged considering the information system renewals and MDM Kalido?

The supporting questions can be divided to three categories of current state, desired state and the future state. First two questions, questions one and two, answer for company's current state and descript the current situation with product data. Third question relates to desired state for renewed information systems, i.e. what kind of solution the company would desire and what is considered the most important part of product data management. Forth question relates directly to future state of the company and how it is going to adapt to information systems renewals with the restrictions from parent company.

1.3. Research objectives and research focus

This research is done for company named Imperial Tobacco Finland Oy (*later ITF*), which is a subsidiary of an international Imperial Tobacco Group PLC (*later ITG*). Objective of the study is to map out ITF's product data's whereabouts and properties. Based on this product data mapping a rudimentary depiction of ITF's data architecture is done. With the current data architecture as a starting point, research's next step is to find out ITF's ideal data architecture from product data perspective in a way that the upcoming information system renewals, those that ITF is aware of and those that might occur in the future, are taken into consideration. With the knowledge of the current, desired and future states a theoretical model for holistic and efficient product data management is introduced as a result of this study.

Research focuses only on the product data therefore excluding the customer and supplier data. To be more precise, the focus on product data is on the master data of product data with the limitation of excluding business processes, data ownership and data types. As the business processes are excluded the IT-perspective is emphasized. Noteworthy point is to acknowledge the fact, that though the IT-perspective is emphasized, some non-IT related issues may rise. Therefore, these non-IT issues are included if they are seen necessary for the objective of the research. Although the master data scope is only limited to product data the other parts of the concept of master data, i.e. business processes and data ownership had to be excluded while research would have expanded because the IT-perspective of the study. The exclusion of data types – as in integer (long/short), string, character, etc. – is done due to the scattered product data within different systems and above all the research would go to detailed level which does not serve the meanings of this research. These limitations and exclusions are made to narrow the scope of the research so that the research itself would fulfill its expectations for ITF within given resources, especially the time horizon.

1.4. Research philosophy

It is stated by von Wright (1970, p.1) that there exists two important aspects in scientific research, one concerning finding the facts and representing those facts and the other concerning formation of the hypotheses and theories. The former is usually called descriptive science and the latter is usually called theoretical science. In addition to these aspects, there are different perspectives on the philosophy of theory formation, interpretation and expounding. These differences of approaches to the scientific research are called philosophies of science. (von Wright 1970, p. 2). Philosophy of science or research philosophy "*relates to the development of knowledge and the nature of that knowledge*" (Saunders et al. 2009, p. 107). Continued with Saunders et al.'s (2009, pp. 108-109) definition of research philosophy, the selected philosophy for the research guides the way how the researcher views the world and the studied phenomena in this context. The selection of particular philosophy influences to research strategies and methods i.e. the way the research is structured and performed. There is no absolute truth of the best philosophy because the philosophy is linked to the research problem and to the research questions, for which the research attempts to answer. (Saunders et al. 2009, pp. 108-109)

According to von Wright (1970, pp. 2-3) the two main philosophies of science are *positivism* and *hermeneutics*, which are each other's opposites. The synonym for hermeneutics is recognized as *idealism*, though von Wright himself avoids usage of this term because its narrow explanation compared to the explanation of hermeneutics. Carson et al. (2001, pp. 6-7) present two research philosophies: positivism as mentioned earlier by von Wright and *interpretivism*. They state that the interpretivism is opposite of positivism and is based on "understanding what is happening in a given context", which sounds similar to hermeneutics presented by von Wright. This contradiction exists due to Carson et al.'s (2001, p. 7) statement, that the term interpretivism is "*derived from the Greek hermeneuein, to interpret*", so instead of using the term hermeneutics, term interpretivism is used (see also Eriksson & Kovalainen 2008, p. 15). Also Eriksson & Kovalainen (2008) and Saunders et al. (2009) speak of interpretivism so therefore and now on, the term interpretivism is used instead of hermeneutics.

The main characteristics of positivism are the consilience of the method, methodological optimum and causal relations. (von Wright 1970, pp. 2-4). The consilience of the method refers to usage of one method that is suitable for all scientific researches despite the object of research. The methodological optimum refers to usage of statistical and mathematical methods through which the optimum level in scientific research can be measured. The causal relations refer to explanation of single causal relationships which are based on generally agreed theories and hypotheses. (von Wright 1970, pp. 2-4; Carson et al. 2001, pp. 6-7).

In positivist philosophy the researcher acts objectively, i.e. he nor she does not affect or is affected by the researched topic (Remenyi et al. 1998, p. 33, according to Saunders et al. 2009, p. 116). As mentioned by von Wright the positivism's use of statistical and mathematical methods refers to quantitative data and quantified observations which give the positivist philosophy a quantitative character. (Saunders et al. 2008, p. 116)

Interpretivism is the opposite of positivistic research philosophy and is more multiform and more heterogeneous than positivism. Interpretivism rejects the idea of the consilience of the method and methodological optimum which are the key concepts of positivism. As the opposite philosophy of positivism, interpretivism focuses on individual and unique features and characteristics which are related to object of research. More commonly said in layman's words, the profound meaning of interpretivism is to understand and to interpret the studied phenomena. (von Wright 1970, pp. 3-4; 25-26). Saunders et al. (2009, pp. 115-116) pinpoint the same crucial properties of interpretivism as the von Wright (1970) and Carson et al. (2001), and emphasize the importance of ignoring the theoretical law-like generalities because the surrounding world and especially the business aspects are far too complex to be generalized. Eriksson & Kovalainen (2008, p. 14) highlights the fact that within research philosophy of interpretivism there may exist different interpretations of the same data because the interpretations are tightly linked to the interpreter, usually the researcher, himself or herself.

As the positivism and interpretivism being the top hierarchical research philosophies, there exist many other philosophies which are based on these dominant philosophies. Saunders et al. (2009, p. 108) have introduced *realism* and *pragmatism*. *Realism* is based on sensations about surrounding world, the reality, which is observed as truth. The realism is divided up to two subcategories: *direct realism*, where our senses and observations illustrate the reality, and *critical realism*, where those sensations and observations are not the illustrations of the reality, only images of the reality because our senses mislead us. *Pragmatism* emphasizes the importance of research question as in choosing the research design, different combinations may be chosen, and i.e. research may have positivist and interpretive characteristics for epistemological, ontological and axiological choices. (Saunders et al. 2009, pp. 109; 114-115; 119).

In addition to Saunder et al.'s research philosophies, Eriksson & Kovalainen (2008, p. 12) and Carson et al. (2001, p. 10) present eight philosophies in addition to above-mentioned: *postpositivism, constructionism, postmodernism, poststructualism, critical theory, humanism, natural inquiry* and *phenomenology*. Also the University of Jyväskylä presents several research philosophies on its web page; some that are not presented by the previous authors (see University of Jyväskylä 2012).

1.5. Evaluation and comparison of research philosophies

Though we have various research philosophies there has to be methods or measurements for evaluation of the different philosophies so the appropriate research philosophy can be chosen for the research. For this dilemma following concepts of ontology, epistemology and axiology are introduced. Also closely relevant concept is research paradigm.

Ontology refers to researcher's view of the reality and how reality functions. Saunders et al. (2009, p. 110-111) present two aspects of ontology: *objectivism* and *subjectivism*. Objectivism depicts reality as entities and phenomenon which are not affected by each other. Subjectivism depicts reality as entities and phenomenon which are affecting each other. Epistemology is concerned with knowledge that will be acceptable in the particular area of research. Axiology depicts researcher's view on the values that will effect on the research. Research paradigm specifies the chosen philosophical choice and reveals the way how the researched phenomenon is interpreted in a way of understanding and explanation. (Saunders et al. 2009, pp. 110-121).

As the research paradigms are the way to study certain phenomena, Saunders et al. (2009, pp. 181-121) present four paradigms: *functionalist*, *interpretive*, *radical humanist* and *radical structuralist*. In functionalist paradigm, researcher tries to find out rational explanations and causal effects to the studied problem, which give the basis for recommendations or rational guidelines for avoiding those problems. In interpretive paradigm, the researcher is concerned about understanding the studied problem and its environment. As the functionalist paradigm is based on rationalities, the interpretive paradigm emphasizes the importance of the irrational events, which give more profound insight to the studied problem. Radical humanist paradigm is related to the idea of changing the current organizational factors, e.g. the methods of work. Radical structuralist paradigm is related to the idea of change that is related to organizational relationships and hierarchies. (Saunders et al. 2009, pp. 120-121)

With these concepts of ontology, epistemology and axiology and with the knowledge of research problem, objectives and limitations, the relevant research philosophy for this research can be chosen. Though there are various research philosophies available, it is more meaningful to do the assessment between two main philosophies, positivism and interpretivism. Ontologically the research is subjective as the researcher is part of the studied organization and also the research setting consists of social factors that may change over time.

As the emphasis of this research is to map out the current situation in an organization, the key point is to understand the current situation and to make interpretations of it. Therefore the epistemological choice would be on interpretivism. The question of axiology is more challenging as it is concerned with researcher's own values. In this case, the researcher is part of the organization and cannot be separated as mentioned earlier, leading to the subjective view and interpretivism.

With interpretivism as a research philosophy the interpretive paradigm is most likely to be adopted. The radical humanist and structuralist paradigms are effective for organizational study though on this case they do not bring any additional value to the research. Interpretive paradigm's concern of understanding the organizational settings is convenient for this research but the functionalist paradigm has to be acknowledged too.

One of the objectives of this research is to create a holistic PDM-model for organization's use. This relates significantly to the functionalist paradigm. When choosing from two possible paradigms, the interpretive one would be the dominant, though the existence of functionalist paradigm cannot be denied. Therefore this research adapts the combination of both paradigms, emphasizing the interpretive paradigm.

1.6. Research strategy and approach

1.6.1. Research approach

Research philosophy answers to the question of what is considered as correct knowledge and its nature and the research paradigm describes how that knowledge will be approached and how it will be exploited. Research strategy refers to ways and techniques which are used to perform the desired research. Saunders et al. (2009, p. 141) state that the choices for research strategy are based on research questions, objectives, resources e.g. time and philosophical choices. As the research philosophy and paradigm are the upper level, abstract like concepts, the research strategy relates to the practical performance of the research, i.e. how the researcher is going to find answers to the research questions, how the research objectives are fulfilled and how the resources are allocated properly (Ghauri & Grønhaug 2005, p. 56).

Before presentation of different research strategies and comparison, debate between quantitative and qualitative research and purposes of research are discussed briefly. Firstly, there is the juxtaposition of quantitative and qualitative research. The choice between quantitative and qualitative depends upon the requirements of the data that is needed for the research and thus relates to the research problem and objectives (Ghauri & Grønhaug 2005, pp. 108-109). In quantitative research the data has certain characteristics such as units of measurement. Therefore the data itself can be processed by statistical or computational methods and results can be presented in statistical or graphical form. (Ghauri & Grønhaug 2005, p. 108; Saunders et al. 2009, p. 414). Conversely the qualitative data has no units of measurements, though the units may be developed but the process can be quite complex and time consuming. The qualitative data can be quantified for quantitative analysis as stated before but mostly the qualitative data requires understanding and interpretation for data analysis because of its unstructured and complex nature, e.g. usage of open ended questions. (Ghauri & Grønhaug 2005, pp. 109-110; Saunders et al. 2009, pp. 480-482)

Secondly, the purpose of the research is related to the structure of the research problem, which can be unstructured or structured and the purposes are distinguished between *exploratory*, *descriptive* and *causal* research (Ghauri & Grønhaug 2005, p. 58). Saunders et al. (2009, pp. 139-140) also present the exploratory and descriptive research purposes but instead of causal research they have introduced *explanatory* research. Exploratory research is suitable for unstructured problems as the problem is vague and it needs clarification. Descriptive research is related to structured problems where the emphasis is to describe the current situation or phenomena, which is seen as an extension or a precursor for exploratory research. Causal research relates to structured problems as well but it focuses on cause-and-effect like problems where key point is to find out the causes that result or lead to different effects. Explanatory research tries to find relationships between studied variables, which on the other hand sounds quite similar to the Ghauri's & Grønhaug's (2005) causal research. (Ghauri & Grønhaug 2005, pp. 57-59; Saunders et al. 2009, pp. 138-140)

1.6.2. Research strategy

Noteworthy issue is the distinction between research strategy and research methods. In this research and thesis, research strategy, e.g. experiment, survey, case study, etc., defines the ways of performance of the research including the conceptual idea of data sources, and the research methods define the tools and techniques, e.g. interviews, observation, etc., for data collection and analysis (Saunders et al. 2009, pp. 141-154).

Research methods are introduced in chapter seven, just before the results of the empirical research. Saunders et al. (2009, p. 141) highlight that there are no superior strategies which would be better over another strategies and nor there are any inferior strategies, some are more suitable then others, depending on the given situation. Saunders et al. (2009, p. 141) present seven strategies for research strategy:

1. Experiment
2. Survey
3. Case study
4. Action research
5. Grounded theory
6. Ethnography
7. Archival research

Saunders et al. (2009, pp. 141-150) describe each strategy on detailed level which eases the evaluation of proper research strategy. Though, the objective of this thesis is not to introduce and compare different strategies on detailed level. Therefore, the case study is the research strategy for this thesis as Robson (2002, p. 178) describes it as "*a strategy for doing research which involves an empirical investigation of a particular contemporary phenomenon within its real life context using multiple sources of evidence*" (according to Saunders et al. 2009, pp. 145-146).

This definition describes the research arrangement opportunely as the research focuses on a case company, its working environment and real time events as in the context and contemporary. Saunders et al. (2009, p. 146) continue, that the case study is relevant choice for getting answers to ‘what’ and ‘how’ questions, as seen in research questions in the chapter 1.3. Though, they remind that the ‘what’ and ‘how’ questions have slight tendency towards survey strategy, which brings out the dilemma of choosing the correct research strategy.

Besides Saunders et al. (2009), Ghauri & Grønhaug (2005) have different presentation of research strategies as shown in the figure 1.1. In figure 1.1, adapted from Ghauri & Grønhaug (2005), they present different research strategies and methods, and divide them between quantitative and qualitative choices.

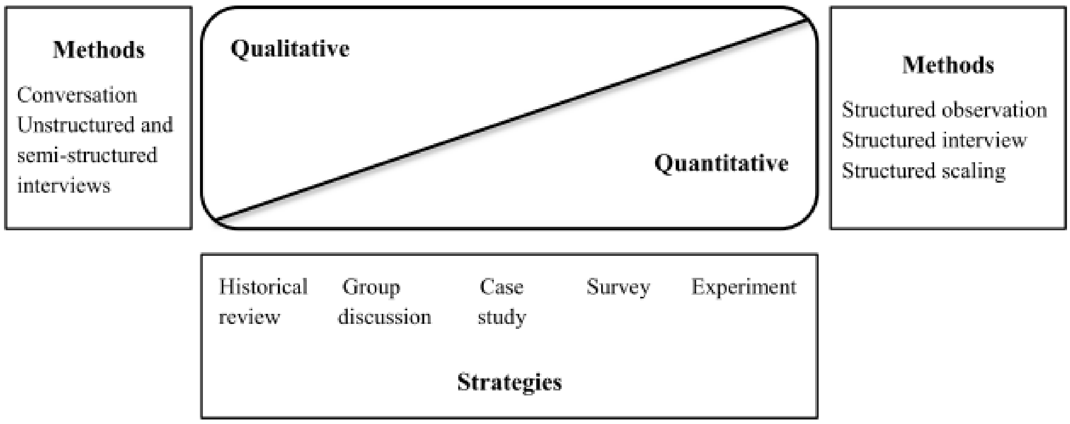


Figure 1.1 *Research methods and strategies from qualitative and quantitative perspectives (adapted from Ghauri & Grønhaug 2005, p. 113).*

There are distinct research strategies in figure 1.1 and every strategy contains qualitative and quantitative features, some more over another. This holds true as studied problem may be quantified but still it needs qualitative perspective for holistic assessment. As the research strategy for this thesis, the case study contains equally quantitative and qualitative features which offer more opportunities for the research performance.

Saunders et al. (2009, pp. 146-147) and Ghauri & Grønhaug (2005, pp. 119-120) refer to Yin’s (1994/2003, different editions of the same source) classification of case studies into two different dimensions: single case versus multiple cases, and holistic case versus embedded case. First dimension, the number of cases refers to the amount of organizations under inspection. Second dimension, the scale of the study refers to units involved in the study, is it an organization as a whole or as different departments. In this thesis, the choice is holistic single case as there is only one company related to the study and there is no distinction between different units or departments. Though the case company is a subsidiary of a global parent company, it is still considered as a single case. (Yin 1994, p. 46, according to Ghauri & Grønhaug 2005, p. 120)

1.7. Structure of the research

So far concepts of research philosophy, paradigms, approach and strategy have been introduced, and methods and data analyzing have been mentioned and they will be introduced in chapter 7 as mentioned before. These concepts make the basis of this research and for upcoming results and their analysis. In figure 1.2 is depicted the research design pyramid.

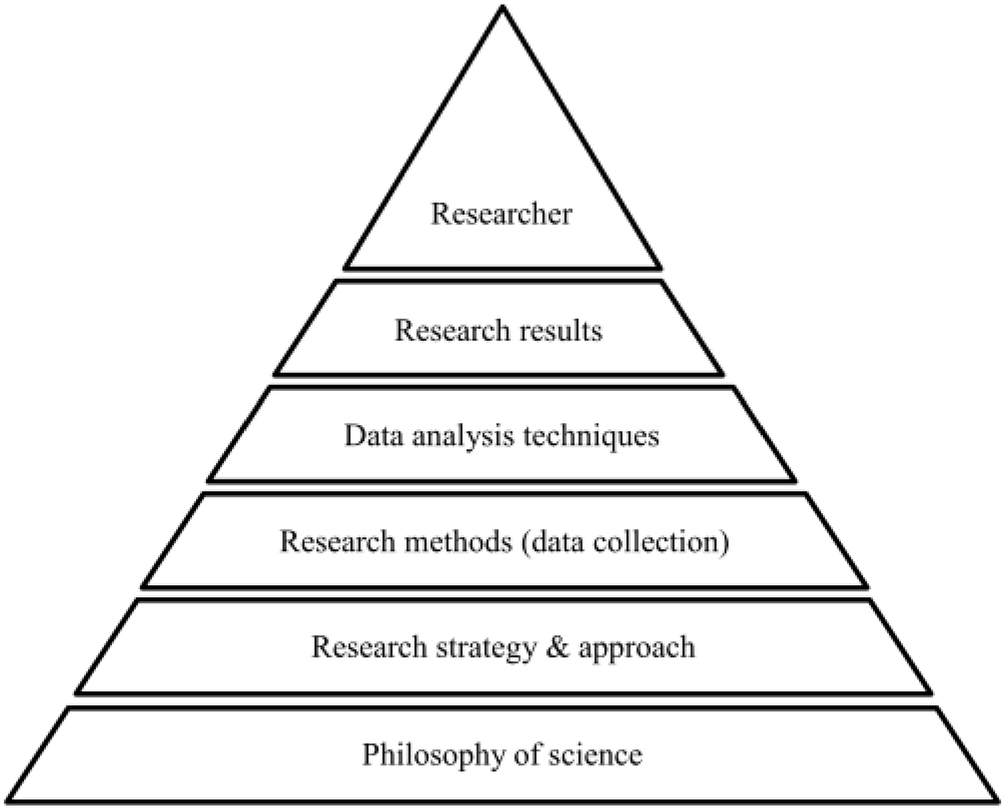


Figure 1.2 *Research design pyramid.*

The key concepts for scientific research are presented in figure 1.2. The research has its foundation in philosophy of science and research paradigm. On top of that is research approach and strategy, and above them is research methods considering the data collection and data analysis. On higher level are the results of the research and uppermost is the researcher himself or herself.

This research is based on interpretivism as philosophy of science and it uses combination of interpretive and functionalist paradigms, emphasizing the interpretive perspective. Research strategy is holistic single case study and by nature it is qualitative research with portions of quantitative aspect. The purpose of the research is quite more complicated to define, as its objectives are so multifold.

On a preliminary state the purpose is to be descriptive and it serves as a basis for exploratory study. On the other hand, the study also focuses on the product data and its whereabouts and content on different information systems and the data architecture itself. That requires explaining and figuring out any relationships between different information systems. Therefore it could be said that the purpose of the research is focused on descriptive and exploratory perspectives with features of explanatory perspective.

About the structure of this thesis, first chapter has described the scientific choices and perspectives for scientific research and the research objectives, limitations, problem and questions. Chapters two and three deal with theoretical background to supplement the empirical findings, from which conclusions and recommendations are made. Chapter four describes the case company and its working environment, and presents the research methods for data collection and data analysis. Chapter five presents the results of the empirical section of the research. Chapter six focuses on discussion about the results and compares those findings to theoretical background. Chapter seven draws final conclusions and recommendations for the case company and the chapter contains assessment of the research itself.

2. KNOWLEDGE MANAGEMENT AND INFORMATION SYSTEMS

Chapter two focuses on knowledge management and information systems. These topics are dealt within chapter two as both provide support for each other. First, concept of knowledge management is discussed in chapter 2.1, which is divided into subcategories concerning definition of knowledge management, process of knowledge creation and framework for knowledge management. Second, concept of information systems is discussed in chapter 2.2, which is divided into subcategories concerning the lifespan of information systems, renewal process of information systems, information systems architecture and data migration.

2.1. Key concepts of knowledge management

2.1.1. What is knowledge?

Before handling the concept of knowledge management, it should be clear that what is meant with the term knowledge. Nonaka & Takeuchi (1995, p. 21) uses Plato's – the well-known philosopher – definition of knowledge, that knowledge is "justified true belief". With this definition knowledge is something that is considered true, until proven otherwise. This is philosophical approach to the definition of knowledge and there exist other approaches as well, more common wise definition for knowledge is presented by Sydänmaanlakka (2004) as seen in the figure 2.1.

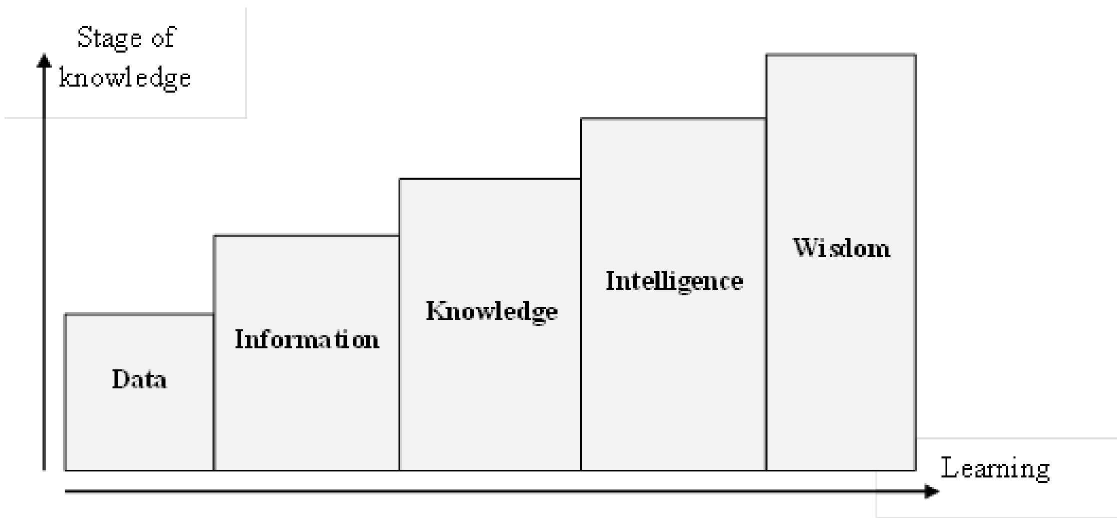


Figure 2.1 *Stages of the knowledge (adapted from Sydänmaanlakka 2004, p. 192)*

In figure 2.1 are the stages of knowledge by Sydänmaanlakka (2004), where the knowledge is divided into five distinct stages from data to wisdom. As we go up to the higher level of knowledge we are simultaneously learning at the same time, i.e. without learning the knowledge will not develop. On the lowest stage is the *data*, which creates the basis for information, which is upper stage of knowledge. Data can be e.g. numbers, text or combination of both, such as weather temperatures. On the next stage is the *information*, which is a set of data that is modified to a relevant and meaningful entity, which informs the interpreter of that knowledge. And by the Liebowitz and Megbolugbe (2003, p. 189), information is a set of data patterned according to certain heuristics. In other words, information contains a message or statement such as “it is cold in the winter, and it is hot in the summer”, alike in the weather temperature example. The data only contains plain numbers like a spreadsheet full of weather temperatures, for a certain amount of time. By examining this spreadsheet, we can develop a statement that: “the temperature is colder in the winter that in summer” which can be reasoned from the data in the spreadsheet. (Sydänmaanlakka 2004, pp. 192-193)

The third stage is *knowledge*, which by itself is very ambiguous term as it is part of stages of knowledge. Due to the ambiguity of term of knowledge, knowledge is more complex and extensive concept than data or information. Knowledge develops from information through learning and it is something that individuals create with their thinking processes. Yet again the temperature example, e.g. it is knowledge that the individual knows to dress in warm clothes during winter time. Next there is the stage of *intelligence*, which is the ability to use right knowledge in the right time. It could be also described as a social intelligence, where individual’s readiness and interaction skills are emphasized. On the highest stage on knowledge is the *wisdom*. Wisdom is about individual’s personal values, moral, beliefs and experiences, which is the result of individual learning process through the lifetime. (Sydänmaanlakka 2004, pp. 194-196)

In addition to Sydänmaanlakka’s (2004) stages of knowledge, Thierauf (2001) had a similar categorization of the stages of knowledge, though he had six stages instead of five. Thierauf’s (2001) six stages are: *data*, *information*, *knowledge*, *intelligence*, *wisdom* and *truth*. As it can be seen, the first five stages are identical (see Thierauf 2001, pp. 7-12, for comparison) and Thierauf’s presentation adds only one new stage, called *truth*. The truth is the highest point of understanding; it is composition of different points of wisdom. (Thierauf 2001, pp. 11-12) At this point it may sound vague and frustrating to compare between different stages of knowledge. Sydänmaanlakka (2004, p. 196) states that it should be efficient to concentrate to the data, information and knowledge perspectives when dealing with the concept of knowledge management. Also, Sydänmaanlakka (2004, p. 196) emphasizes that the intelligence and wisdom should not be forgotten, though they are more relevant for the perspectives of individual.

Although there are different stages of knowledge by Sydänmaanlakka (2004) and Thierauf (2001), there is also another classification for knowledge which was introduced by Michael Polanyi in the 1950's (see Polanyi, M. 1958. Personal Knowledge, for further detail). Though this classification was done several decades ago, it was made famous by the Nonaka & Takeuchi (1995). (Sydänmaanlakka 2004, p. 197; Nonaka & Takeuchi 1995, p. 59) This classification divides knowledge to *tacit knowledge* and *explicit knowledge*. The tacit knowledge is personal, within the individual and it depends on the context, which on the other hand makes it hard to write down for distribution. The explicit knowledge is the opposite of tacit knowledge as it is has formal style of presentation, e.g. written form and it is systematic in nature e.g. instructions for furniture. (Nonaka & Takeuchi 1995, p. 59) Sydänmaanlakka (2004, p. 197) adds that the tacit knowledge is difficult to transfer to other individuals and the tacit knowledge consists of actions that we are capable to do but are difficult to describe, e.g. riding bicycle. And the explicit knowledge is objective and transferrable, such as documents or blue prints.

As it is presented, the knowledge itself is a multifold concept which can be defined in various ways. The stages of knowledge give an excellent way to compare and assess the types of knowledge. Distinction between data and information, and between information and knowledge are most evident whereas intelligence, wisdom and truth are more abstract and more difficult to describe. And to make matters more complicated, there is the classification of tacit knowledge and explicit knowledge. Sydänmaanlakka (2004, p. 197) emphasizes that distinction between tacit and explicit knowledge is usually neglected and people are unaware of the amount of tacit knowledge compared to explicit knowledge. So far the concept of knowledge is introduced and revealed, thus the process and methods of knowledge creation has to be examined.

2.1.2. Knowledge creation

Previous chapter defined the concept of knowledge, this chapter defines the processes and methods for knowledge creation. Nonaka & Takeuchi (1995, p. 56) describe an example of organizational innovation, where the action of innovation is not merely processing the current information rather it is process of creating new knowledge and information in order to redefine the problems and solutions. By so, the organization re-creates its own environment and may gain competitive advantage against other organizations. Respectively, Sydänmaanlakka (2004, pp. 179-180) acknowledges the importance of knowledge to the organizations and remarks that organizations may not know what they know or what they should know. Sydänmaanlakka (2004, pp. 187-189) sees the knowledge creation as a one sub-process of knowledge management, whereas the other four sub-processes are *procurement*, *warehousing*, *distribution* and *application*. These sub-processes will be covered in the next chapter and here we will concentrate on knowledge creation and start off with dimension of knowledge creation, as presented by Nonaka & Takeuchi (1995) in figure 2.2.

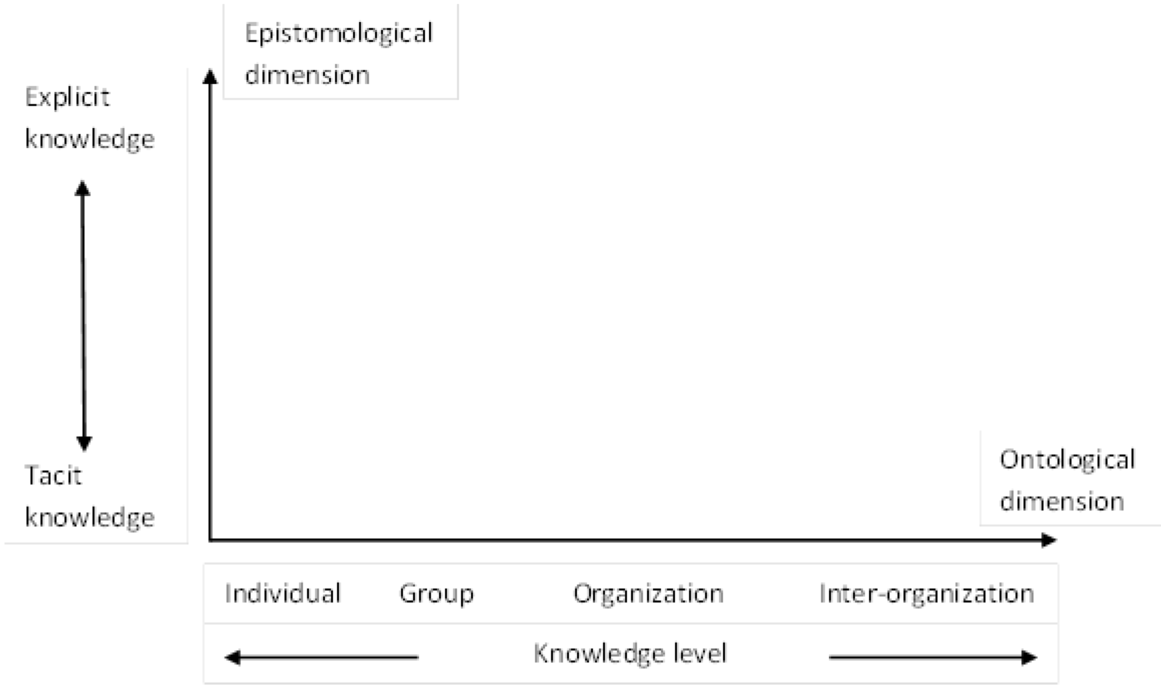


***Figure 2.2** Dimensions of knowledge creation (adapted from Nonaka & Takeuchi 1995, p. 57)*

In figure 2.2 Nonaka & Takeuchi (1995) have presented two dimensions for knowledge creation. The dimensions are epistemological and ontological, where the epistemological dimension refers to the theory of knowledge and the ontological dimension is related to the knowledge creating entities. Theory of knowledge is divided to tacit knowledge and explicit knowledge and the knowledge creating entities vary from individual to inter-organization. (Nonaka & Takeuchi 1995, pp. 56-57) In addition with Nonaka's & Takeuchi's (1995) dimensions of knowledge creation, the stages of knowledge – by Sydänmaanlakka (2004) and Thierauf (2001) – should be included to the knowledge creations process. Thus, the knowledge creation alters from two dimensional processes to three dimensional processes.

As it comes to knowledge creation, the process starts on the individual level when the explicit and tacit knowledge are combined and moves on the higher levels of ontology, towards inter-organizational level. Inclusion of stages of knowledge (see figure 2.1) adds third dimension, thus the knowledge creation begins as described above, but starting from the lowest levels and ascending upwards, from data towards wisdom (truth)[1]. The creation process starts within individual level as organization itself is not capable of creating knowledge. From the individual level, the knowledge creation spirals through from explicit to tacit and from tacit to explicit knowledge, and converting the knowledge.

[1] Stages of knowledge by Sydänmaanlakka (2004) begins from data and ends to wisdom, and the corresponding presentation of stages of knowledge, by Thierauf (2001), presents one additional stage, truth, that is the highest stage.

In parallel with these knowledge conversions, the ontological level is elevated and individuals' knowledge becomes group's knowledge and so on, and simultaneously the stage of knowledge changes as the data transforms to information and further to the knowledge. As there is distinction between explicit knowledge and tacit knowledge, then there are only four possible knowledge conversion processes; from tacit to tacit, from tacit to explicit, from explicit to tacit, and from explicit to explicit. These knowledge conversions are shown in the figure 2.3 and those conversions are usually referred as SECI-model. (Nonaka & Takeuchi 1995, pp. 57-59)

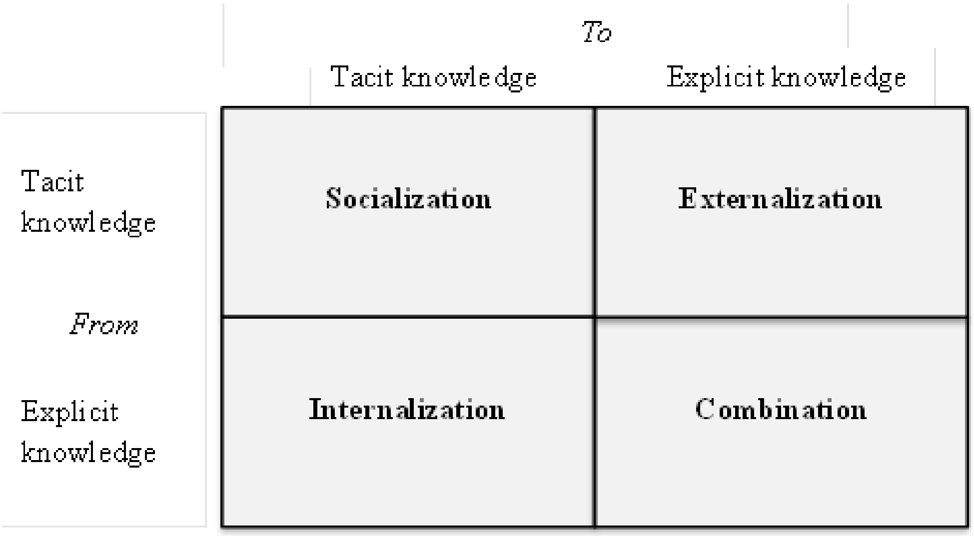


Figure 2.3 *Knowledge conversions a.k.a. SECI model (adapted from Nonaka & Takeuchi 1995, p. 62)*

In figure 2.3 are the four knowledge conversions, which are named as: *socialization*, *externalization*, *combination* and *internalization*. The reference to SECI-model comes from the first letters of the names of knowledge conversions. Figure 2.3 also shows the type of knowledge conversions, as from where to what. Noteworthy point with Nonaka's & Takeuchi's (1995) knowledge conversions is that they do not take stand on stages of knowledge, presented by e.g. Thierauf (2001). For simplification, the conversion types and processes are summarized in table 1.

Table 1. *Summary of knowledge conversion types.*

Conversion	**Socialization**	**Externalization**	**Combination**	**Internalization**
From	Tacit	Tacit	Explicit	Explicit
To	Tacit	Explicit	Explicit	Tacit

Socialization is a conversion from tacit knowledge to tacit knowledge, as seen in table 1. Basically socialization is about sharing experiences and learning through observation, imitation and practice. Such as apprentice-master situation where apprentice learns by observing master's work and functioning. *Externalization* is a conversion from tacit knowledge to explicit knowledge. This conversion is quite complicated as the main idea is to transform individual's personal knowledge into explicit form, as a written document. Therefore the conversion's results are models and concepts that try to describe the tacit knowledge, e.g. instructions for riding a bicycle. (Nonaka & Takeuchi 1995, pp. 62-64)

Combination is a conversion from explicit knowledge to explicit knowledge. As the explicit knowledge is transferrable and possesses a formal style of presentation, the combination conversion is about combining and rearranging the explicit knowledge in a novel way. E.g. monthly reports are produced by combining different documents and data sets to present existing explicit knowledge in a distinct way. *Internalization* is a conversion from explicit knowledge to tacit knowledge. This conversion can be described as learning through doing, as an individual internalizes knowledge which he or she has used for certain action. By reading instructions, e.g. for building a birdhouse, and working accordingly to them, the individual learns and internalizes the working process and adapts the knowledge as his or hers personal knowledge. (Nonaka & Takeuchi 1995, pp. 67-70)

Finally, the four knowledge conversion types are presented which depict the interaction in the epistemological axis in a figure 2.2. Also the knowledge creation involves the ontological axis, which includes the knowledge creating entities. Nonaka & Takeuchi (1995, p. 70) emphasizes that the use of one knowledge conversion solely is not sufficient for organization's knowledge creation purposes, rather the knowledge creation process is continuous and dynamic process, combining all four conversion types and utilizing all the knowledge creating entities. Within this continuous and dynamic process of knowledge creation, the addition of stages of knowledge to the Nonaka's & Takeuchi's (1995) knowledge creation process helps the follow-up of knowledge creation. That is, as the knowledge is converted back and forth between explicit and tacit knowledge, and the knowledge creating entities are utilized, it is crucial to recognize that on which stage the process is executed. If the certain set of knowledge has reached the stage of wisdom, is it plausible to continue with that set on that stage, or should the focus of knowledge creation be focused on other areas, e.g. knowledge that is on the lower stages?

2.1.3. Knowledge management framework

Sydänmaanlakka (2004, p. 187) states that knowledge management is a process where knowledge is *created*, *procured*, *warehoused*, *distributed* and *applied*. These supporting processes endorse the knowledge conversions and ontological perspective of knowledge creation, i.e. individual's knowledge will become group's knowledge (Sydänmaanlakka 2004, p. 187). Kamara et al. (2002, p. 205) state that knowledge management is seen as a tool, which enables organization to define its need for innovations and improvement of business. Sydänmaanlakka (2004, p. 186) emphasizes, that the most founding idea of knowledge management is to provide sufficient knowledge for the decision making processes, which on the other hand improves business making as seen as an objective of knowledge management by Kamara et al. (2002). The knowledge should be significant and contain meaning in order for applying the knowledge in practice. In other words, though the knowledge can be old or novel, it has to be useful for organizational purposes; otherwise its utility value is rather low. (Sydänmaanlakka 2004, p. 186)

Sydänmaanlakka (2004, p. 180) claims that the definition of knowledge management is vague and the definition depends on the definer, so it is useful to define the concept of knowledge management through a framework which models the reality. Sydänmaanlakka (2004, p. 181) approaches the knowledge management framework from three perspectives: *organizational*, *cultural* and *individual/group*. Organizational perspective includes organization's strategies, objectives, information systems, human resource management and supporting systems for learning. Cultural perspective includes organizational culture, managerial issues, communication, feedback and organizational values. Individual/group perspective consists of ability and willingness to learn, change, apply, and to distribute and receive knowledge. (Sydänmaanlakka 2004, pp. 180-182)

In consensus with Sydänmaanlakka's (2004) perspectives on knowledge management, Kamara et al. (2002, p. 205) state that the knowledge management implementation solely from the IT-perspective is determined to be the opposite of successful, as the organizational and personnel issues has to be considered as well. In addition, Liebowitz's & Megbolugbe (2003, pp. 189-190) share the similar perspective as they state that organizational culture, individual's values and organizations view of the world affects knowledge, and therefore its management. These are the fundamental perspectives for knowledge management framework and functioning in the organization (e.g. Sydänmaanlakka 2004). Although the described functions and activities may sound quite general and vague, the most important issue is to change the way of thinking and adapt our minds to current information society. So, for a clarifying depiction, in the figure 2.4 is a model of achieving benefits from organizational strategy by applying knowledge management functions (Sydänmaanlakka 2004, p. 204).

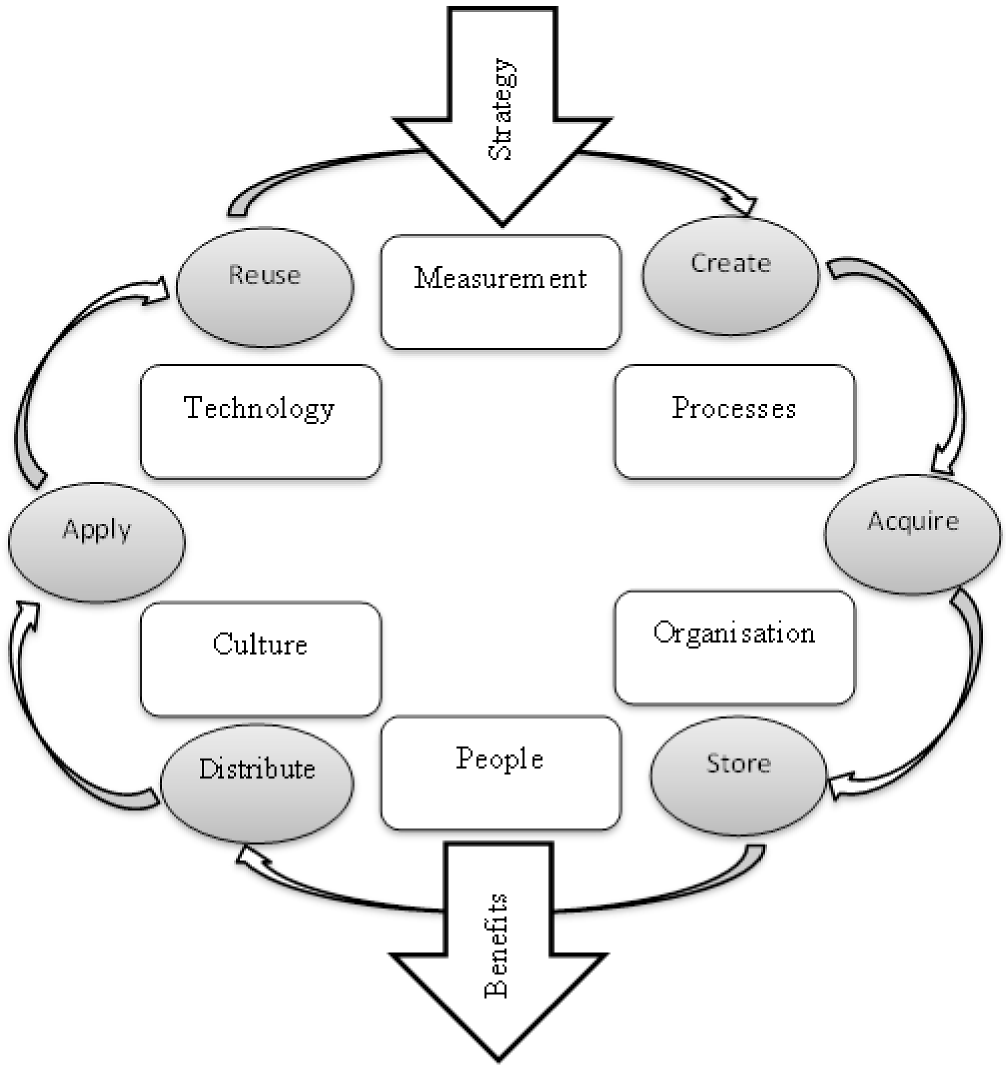


Figure 2.4 *From organizational strategy to organizational benefits by the means of knowledge management (adapted from Sydänmaanlakka 2004, p. 204)*

In a figure 2.4 the process starts with *measurement*, where the current situation with performance and competence are measured. This phase includes measurement of existing knowledge and questions the sufficiency of knowledge and conditions with sub-process functions of knowledge management, i.e. create, acquire, store, distribute, apply and reuse. If there is sufficient knowledge and it is managed properly by creating new knowledge, acquiring it and so on, then the process of knowledge management is performed precisely and brilliantly. (Sydänmaanlakka 2004, pp. 204-205) For scale and perspective, the sub-process of knowledge creation is described in previous chapter (see chapter 2.1.2). As the knowledge creation process was seen as three dimensional process – it combines Nonaka's & Takeuchi's (1995) original two dimensional process enhanced with stages of knowledge, by Sydänmaanlakka (2004) – it requires a lot to succeed. And the knowledge creation is only one of the six sub-processes, therefore the whole process of knowledge management is everything but simple.

If the sufficiency of knowledge is not on a satisfactory level, then the components of knowledge management – *processes, organization, people, culture* and *technology* – should be worked through and revised (*see: figure 2.4*). Some of the following questions may rise and should be noted (adapted from Sydänmaanlakka 2004, p. 205):

- Are we using best practices (*processes*)?
- Does organizational structure support knowledge management and its objectives (*organization*)?
- Are the employees aware of the meaning of knowledge management initiatives (*people*)?
- Is the organizational culture open for changes (*culture*)?
- Do we possess sufficient IT-technology and tools to enable knowledge management functions (*technology*)?

The knowledge management framework in figure 2.4 depicts the crucial components for successful knowledge management which on the other hand aid the fulfillment of the organizational strategy, which yields to organizational benefits. Those crucial components consist of knowledge management sub-processes – from creation to reuse – and the organizational factors – from measurement to technology. When these processes and factors are studied and managed with precision, the organization will gain competitive advantage and flourish. Above all, it should be emphasized that the knowledge management is dynamic and continuous process that should not be seen as a one-time arrangement. (Sydänmaanlakka 2004, pp. 205-207) Similar framework (compared to figure 2.4) is presented by Liebowitz & Megbolugbe (2003, p. 190), which encompasses components including tools, technologies, strategies, measurement and organizational structure for knowledge management.

Nature of knowledge is complex, as it varies from data to wisdom; it can be classified between tacit knowledge and explicit knowledge; and it includes and involves the knowledge creating entities (see figure 2.2). Within these dimensions the knowledge management may sound simple; just categorize the stage of knowledge, make the difference between tacit and explicit knowledge, and determine the knowledge creating entities. But, the knowledge management takes place within continuously developing business environment and involves the organization and its employees. Also there is the distinction between individual and group, and between group and the whole organization. Combine all this with information technology and information systems and you will realize how complex everything can be. Agonizing as it may seem, knowledge can be managed and with pride, all you have to do is plan before you act, and measure and supervise the achievements. (Nonaka & Takeuchi 1995; Thierauf 2001; Kamara et al. 2002; Liebowitz & Megbolugbe 2003; & Sydänmaanlakka 2004)

2.2. Information systems

2.2.1. Lifespan of an information system

The concept of knowledge management relates closely to information systems and their development. Before examining concepts of information systems, the definition of information system is presented by Turban et al. (2008): "An information system (IS) collects, processes, stores, analyzes, and disseminates information for a specific purpose. -- Like any other system, an information system also includes people, procedures, and physical facilities, and it operates within an environment. An information system is not necessarily computerized, although most of them are." (Turban et al. 2008, p. 16) From Turban et al.'s (2008) definition it can be clearly seen that information systems and knowledge management have something in common, such as the processes of collecting, storing and disseminating, and moreover information systems can be applied for knowledge management purposes.

Function of an information system[2] is to enable business for organization in a way it was intended, i.e. CRM (*customer relationship management*) systems are for management of customers. As business evolves, the technologies evolve as well, which lead to better information systems and greater benefit to the organization, e.g. increased functionality and flexibility or impacts on the organizational information system infrastructure (Irani et al. 2003, p. 178). Naturally, information systems have a certain lifecycle or a lifespan, which consists of time from the development of the information system to the point where the information system is removed from the usage (Haikala & Märijärvi 2006, p. 36). To understand the complexity of an information system, the creation and development process of an information system is depicted in figure 2.5.

Figure 2.5 depicts traditional process of an information system development and it is most commonly referred as a waterfall model. The reference to waterfall comes from real life analogy of waterfall where water flows from the top down. So, alike real life waterfall, you move from up to bottom like the water, e.g. the idea of this model is to complete previous stage before proceeding to the next stage. Though it may be possible to move to the opposite direction of the waterfall, it is quite difficult as it sounds. Thus, the information systems development begins with the requirements study stage, descending downwards as the previous stage is completed, and ending with the implementation stage. [3]

---

[2] Though the definition of information systems by Turban et al. (2008) included people, procedures and facilities within certain environment, from here on the information system refers for most of the part to the application or software, such as Microsoft Word. The clarification is done because the use of the term information system is vague and depends on the context.
[3] Though the waterfall model is quite rigid, the main idea is to describe the process of information system development, and there exist other models for similar development but they are excluded from this research as they would not bring any added value to the studied phenomenon.

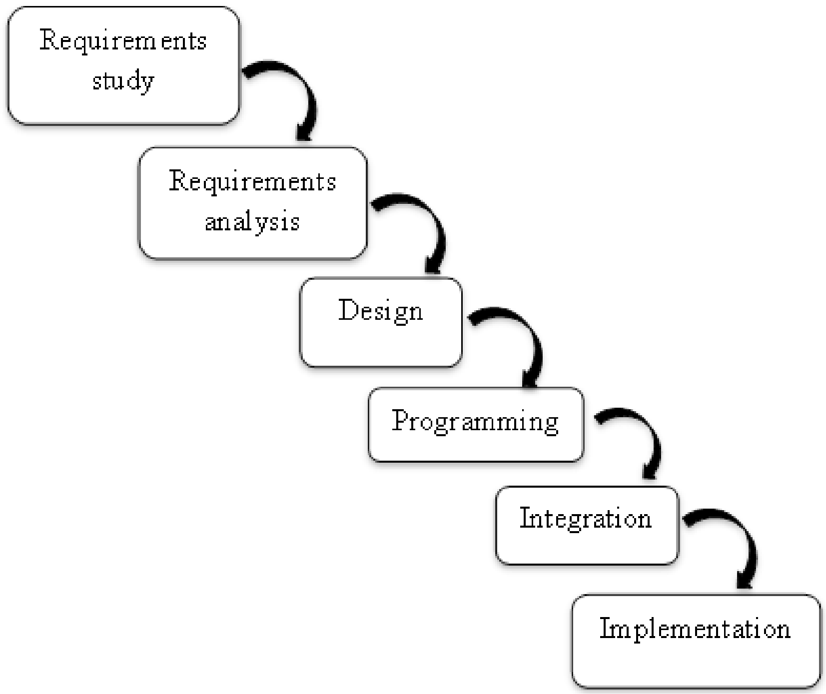


Figure 2.5 *Information system development via waterfall model (adapted from Haikala & Märijärvi 2006, p. 36)*

The waterfall model in figure 2.5 consists of six distinct stages: *requirements study*, *requirements analysis*, *design*, *programming*, *integration* and *implementation*. *Requirements study stage* includes determination of the customer's needs for the information system. In *requirements analysis stage* customer's needs are analyzed and based on the customer's requirements, functional specification is done which encompasses information system's operations and functions, e.g. how to invoicing is done, appearance of the user interface and possible restrictions. In *design stage* the functional specification is transformed to technical requirements, which describe the technologies and architectural choices about how the information system is build. E.g. technical requirements describe that the information system is composed of several modules, and the modules represent a certain function in organization's business, like accounting, marketing, etc. (Haikala & Märijärvi 2006, pp. 38-40)

Programming stage includes all the time and work that begins from writing the code with programming language to the point where the code is compiled without errors. In other words, the information system is produced in a language that is understood by the computer and the information system is functioning, i.e. it can be used. Noteworthy point is that in this point and through the programming and integrations stages, the written code is tested to ensure information system's correct functioning. In *integration stage*, all the separate parts, alias modules, of the information system are integrated together and their functioning is tested. In *implementation* stage the information system is adopted in the organization (customer) and the usage of the information system begins. (Haikala & Märijärvi 2006, pp. 40-41)

After the information system is developed and implemented to the customer's environment, the maintenance function begins. Maintenance includes defect correction but it may also include perfective maintenance which aims to improve functionalities of the information system by altering or adding features. (Haikala & Märijärvi 2006, p. 41) The perfective maintenance is quite similar to the product development or to the further development, where the features and functionalities of the information system are enhanced by the feedback from the customer. (Haikala & Märijärvi 2006, p. 44) Related to the knowledge management, information systems development requires knowledge management in order to execute the development process from start to finish. Also, the development process begins with scarce and undefined requirements, which will improve and specify during the process, alike in knowledge creation, the stage of the knowledge will rise.

As Haikala's & Märijärvi's (2006) development model for information systems is quite technical and moreover related to the information system developer's perspective, it can be used to describe lifespan of an information system. Though, the focus is on the beginning of the lifespan as the waterfall model excludes the usage and removal of the information system. In addition to Haikala's & Märijärvi's (2006) waterfall model, Berghout et al. (2011, p. 756) have taken more financial perspective to information systems and their development.[4] They emphasize the importance of ratio between costs and benefits, which is one factor that should be noticed when discussed about lifespan of an information system. During a development stage, the costs will increase and after implementation, the costs will decrease but within certain time, aging problems, e.g. old technology, etc., will arise and the information system has come to the end of its lifespan. (Berghout et al. 2011, p. 756)

2.2.2. The information system renewal process

Despite the fact that there are efficient information systems, there is still need for information system renewal as the current information systems become obsolete, e.g. aged technology or altered requirements. Information system should be renewed, when it does not fulfill the purpose for which it was designed and developed to do. The reason of renewal may relate to cost factors, lack of compatibility with other information systems or corporate merger, which may lead to compliance issues between two organizations. These are only some of the reasons for renewal and the renewal process is closely linked to the concept of information system's lifespan (see chapter 2.2.1). The renewal process will be reviewed from Berghout et al.'s (2011) financial perspective as well as from Turban et al.'s (2008) process-based perspective.

---

[4] Berghout et al.'s (2011) information systems development model could be adapted to the lifespan concept (from the financial perspective), though it was seen as a preferable option to present their model within information systems renewal topic, in chapter 2.2.2.

A model by Berghout et al. (2011, p. 756) depicts information system renewal process through information system's lifecycle and explores the process via financial perspective. The model consists of five different stages (Berghout et al. (2011, p. 756)):

1. Identification of project proposals
2. Justification of IT proposals
3. Realization of IT projects
4. Exploitation of the information systems
5. Evaluation of IT cost/benefit management

Identification of project proposals includes the decision making process about the IT resources and projects, e.g. information system, which require investments and are required from organization's perspective. This stage also includes the assessment of costs and benefits for every development project, i.e. what benefits the new information system brings and what does it cost. *Justification of IT proposals* stage deals with IT proposals which should be invested and implemented. (Berghout et al. 2011, p. 756)

Realization of IT projects stage relates directly to the information system development project (*see figure 2.5 for revision*), where costs and benefits are managed through measurement and monitoring. *Exploitation of the information systems* stage includes the usage of the information system as well the management of operational costs and benefits of the renewed information system. *Evaluation of IT cost/benefit management* stage tries to answer to the question of success of the IS investment. In other words, are the invested costs worth of the gained benefits, and if not, then what are the main reasons for the failure. Also, this stage includes the efficiency perspective for organization of the IT resources; could it be done more efficiently? (Berghout et al. 2011, p. 756)

Berghout et al.'s (2011) model is quite straight forwarded process starting from the identification of the IT projects which should be realized, then measured and monitored via management practices and ending up with the usage of the information system and its cost-to-benefits evaluation. In contrast to Haikala's & Märijärvi's (2006) model, Berghout et al.'s (2011) model begins with identification for possible IT projects, whereas Haikala's & Märijärvi's (2006) model begins with requirements determination, i.e. the IT-project has already identified earlier. As the iteration of this model is done, the second iteration phase begins, starting with the identification stage. Though, this can be seen as iterative and periodical process, it should be noted that the process can start from different stage than the identification stage. E.g. the evaluation stage is the phase where the possible flaws or defects are observed and these observations serve as an input for the identification stage. (Berghout et al. 2011, pp. 756-757)

In addition to Berghout et al.'s (2011) model, another model by Turban et al. (2008) is presented. Alike the Berghout et al.'s model, Turban et al.'s model consists of five different stages (Turban et al. 2008, pp. 594-596):

1. Planning, identifying, and justifying IT-based systems
2. Creating an IT-architecture
3. Selecting an acquisition option
4. Testing, installing, integration and deploying IT applications
5. Operations, maintenance and updating

Planning, identifying, and justifying IT-based systems stage begins with the need for a new information system, which can be caused due to changes in the business environment. If the needs and requirements for the renewed information system are known, then this stage focuses and deals with compatibility and business suitability issues. *Creating an IT-architecture* stage refers to IT-resources of the organization which are used to deliver the desired result, i.e. information, data and information system functionalities. In this stage it is important to scope out the possibilities – i.e. interfaces, etc. – which are available and plausible for the renewed information system. *Selecting an acquisition option* stage is phase where the choice of renewed IS-solution is made. This stage includes evaluation and assessment of distinct issues and factors between different IS-solutions. The objective is to find out the plausible solution that meets up to the expectations and requirements of the organization. E.g. evaluation can be between the executions like SaaS (*Software-as-a-Service*) versus SOA (*Service Oriented Architecture*). (Turban et al. 2008, pp. 595-596)

Testing, installing, integration and deploying IT applications stage starts with information system's installation and integration to the other information systems. Then the information system's functionalities are tested in the organization's environment and after successful testing the information system is deployed for organization-wide usage. This stage also includes the user training and the change management.[5] In *operations, maintenance and updating* stage the information system is being used and the maintenance function is carried out by the maintenance contract between information system's supplier and the organization. Depending on the information system, the supplier may develop new functions and versions of the delivered information system. (Turban et al. 2008, pp. 596-597)

Turban et al.'s (2008) model is generic and it can be applied to the acquisition of a new information system or to the development of old information system. The model is more process-based than the financially emphasized model of Berghout et al. (2011). As the

---

[5] Noteworthy point is the change management, which should play crucial role in the information system renewal and acquisition, though this entity is excluded from this research due to several limitations.

waterfall model of Haikala & Märijärvi (2006) was more from the IS developers perspective, the model of Turban et al. (2008) is focused on the customer perspective, i.e. the acquirer of the information system.[6] Though, these models of Berghout et al. (2011) and Turban et al. (2008) supplement each other, and based on these models a combined model for information systems renewal is introduced in figure 2.6.

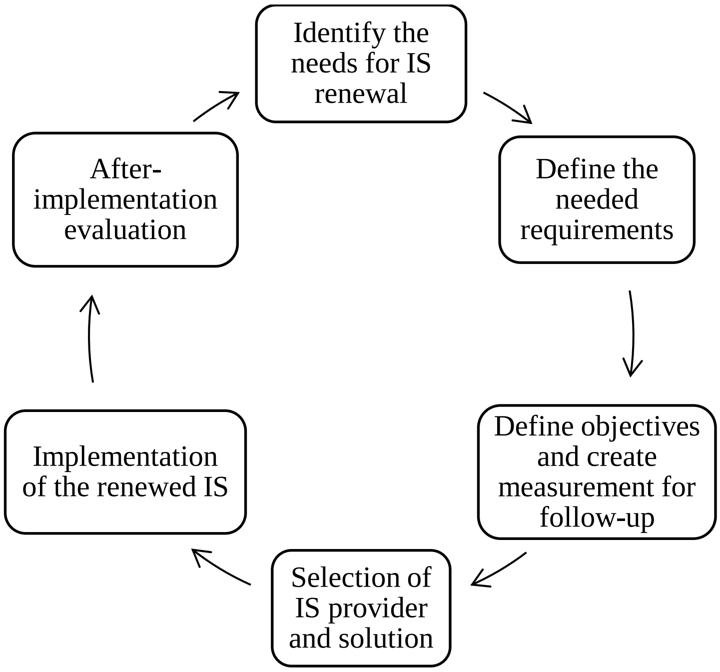


Figure 2.6 *Iterative model for information systems renewal (adapted from Turban et al. 2008; Berghout et al. 2011)*

In the figure 2.6 is an iterative model for information systems renewal and the founding idea is that the process is continuous and emphasizes the fact that the process should not be stopped after the implementation. This is because, when the process is continuous then the organization can respond to the changes far more easier and faster than in the situation where the process starts from void. The stages in this model do not vary much from the presentations of Turban et al. (2008) and Berghout et al. (2011), though it combines the best qualities from both of them and introduces the process as iterative and periodical. The stages are introduced in a question-form in a list below [7](Turban et al. 2008; Berghout et al. 2011):

1. Identify the needs for IS renewal
    - What are the issues for IS renewal?
    - Are the issues related to the IS or working habits?

---

[6] Though the waterfall model was presented within IS lifespan concept, it can be seen as a IS renewal process from the developer perspective. Alike the footnote 4 states that Berghout et al.'s (2011) model could be used for both purposes and same could be stated for Turban et al.s (2008) model.
[7] The list itself does not try to be comprehensive and it is more suggestive in nature than concrete list of executable actions.

2. Define the needed requirements
   - What new do we require and need, when compared to the older IS?
   - Which requirements are essential and which are useful but not essential?
3. Define objectives and create measurement for follow-up
   - What are the milestones within this project?
   - What are the measurements by which the success of this project can be measured?
4. Selection of IS provider and solution
   - Reputation of the provider?
   - How well the provided solution is adaptable to organization's environment?
5. Implementation of the renewed IS
   - How it will be done and does it include to the costs?
   - How the user training is done?
6. After-implementation evaluation
   - Does the renewed IS fulfill the objectives?
   - Are the benefits achieved within given budget?
   - Does the IS have all needed functionalities for organization's business?

The list above tries to wake discussion about information system projects and possible renewals. For example, is the information system renewals right option for gaining business advantage or are the reasons on working habits and methods? And when the renewal process comes relevant, it should be considered that what are the tangible objectives that organization tries to fulfill via renewal process? I.e. the objectives should be measurable in order to follow the success of the renewal process. By working through these stages and acknowledging the possible obstacles that may rise, the information system project should go well, compared to the project where something is done without measurement and proper management.

2.2.3. Information systems architecture

Oxford Reference Online (2013) defines that architecture is "the term given to an organization's information technology platform, structure and process --", whereas the Collins English Dictionary (2013) defines architecture as "the internal organization of a computer's components with particular reference to the way in which data is transmitted" or as "the arrangement of the various devices in a complete computer system or network". On the other hand, the definition of architecture is very common to layman as it comes to buildings and their appearance and structure. Though, the professional association IEEE has defined the term architecture from IT perspective in its standard IEEE 1471-2000: "architecture is the fundamental organization of a system embodied in its components, their relationships to each other, and to the environment, and the principle guiding its design and evolution." (Jonkers et al. 2006, p. 63)

The definition by IEEE standard and the dictionary definitions mean that the architecture describes the holistic entity, its components and their relations to each, such as blueprints. Whit this knowledge of architecture helps us to design and reconstruct the existing structure, e.g. organization's structure. For example, architecture describing the whole organization and its processes, information systems and infrastructure is called as enterprise architecture (Jonkers et al. 2006, p. 64). As there is a need for change regarding certain processes or information systems, then the architecture – in this case enterprise architecture – helps the development process when all components concerning the change can be taken into consideration. Meaning, to where the information system is related and to where its actions affect? (Jonkers et al. 2006, p. 64)

Jonkers et al. (2006, p. 64) state that the enterprise architecture is usually constructed from separate architectural domains, each domain concentrating to its specific purpose, such as technical architecture, process architecture, application architecture, and so on. From the perspective of this research, the architectural examination is focused on the information systems architecture, which encompasses organization's selection of information systems and their relations in between.

Kurbel (2008, p. 97) defines the information systems architecture in same fashion, though he refers to the architecture as a large information system, containing subsystems such as applications, etc. whereas in this research, information system means same as Kurbel's subsystem. Because the reality of information systems architecture is more than plain information systems, also related parties, i.e. suppliers and customers, interfaces, databases and other technical resources are included to the definition of information systems architecture in certain extend. As the term architecture and information systems architecture are defined for this research, a depiction of information systems architecture is presented in figure 2.7.

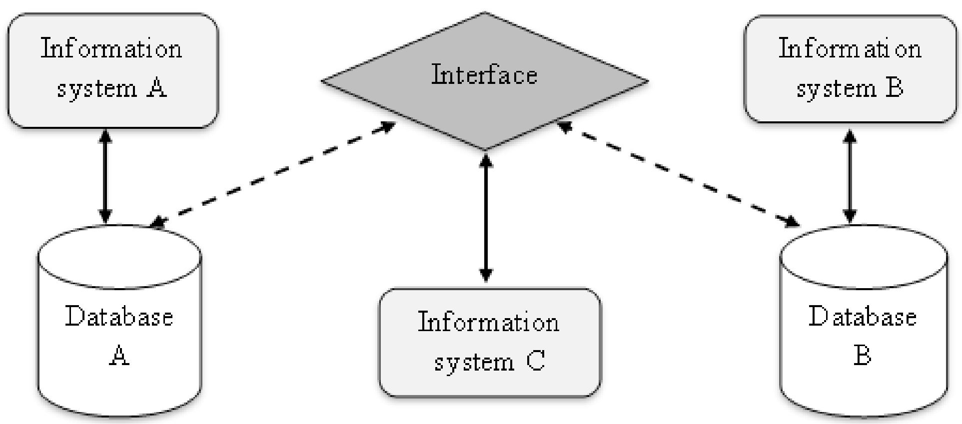


Figure 2.7 *Depiction of information systems architecture (adapted from Jonkers et al. 2006; Kurbel 2008)*

In figure 2.7 is a depiction of information systems architecture which is used for clarification purposes. The information systems architecture is composed of following components: three separate information systems (A, B & C), two separate databases (A and B) and one interface operating between databases A and B, and information system C. Information system A uses database A for its own specific purpose, e.g. customer relationship management. Like information system A, information system B uses database B for its own specific purposes. Finally, information system C is for customer orders and delivery and it retrieves required information for order handling from the databases A and B by using given interface. If one or more of these components are changed or replaced by other solutions, the information systems architecture helps to find out the components which are also affected by this change. E.g. information system C is replaced with information system D, then what should be done with the interface and are the databases A and B sufficient for information system D's requirements?

Depiction of information systems architecture in figure 2.7 is quite simple and plain, though it gives a quick review to the world of information systems architecture. When dealt with changes in information systems architecture, it gets more complicated than replacing system C with system D. Kurbel (2008, pp. 98-105) inspects also architectural patterns of a single information system.[8] Therefore, when speaking of information systems architecture, the information systems must be treated as a single components as well as part of larger entity, i.e. a component that is part of all the information systems in organization (Kurbel 2008). Though the information system architecture is excluded from theoretical examination, it should be noticed that without it, the development of the current architecture can be rather difficult due to lack of knowledge, which on the other hand can be improved with proper knowledge management.

Ardagna et al. (2004, p. 231) have presented a methodology for designing IT architecture which on the other hand helps the development of information systems architecture. The model relates to more detailed and technical information, consisting of IT resources and technologies. The model by Ardagna et al. (2004, p. 231) is presented in figure 2.8.

---

[8] To this point the term information systems architecture has been used in a plural form, so that the term would include organization's all information systems. Though, the term information system architecture, a singular form, is only used on this paragraph as the emphasis is on the *information systems* architecture.

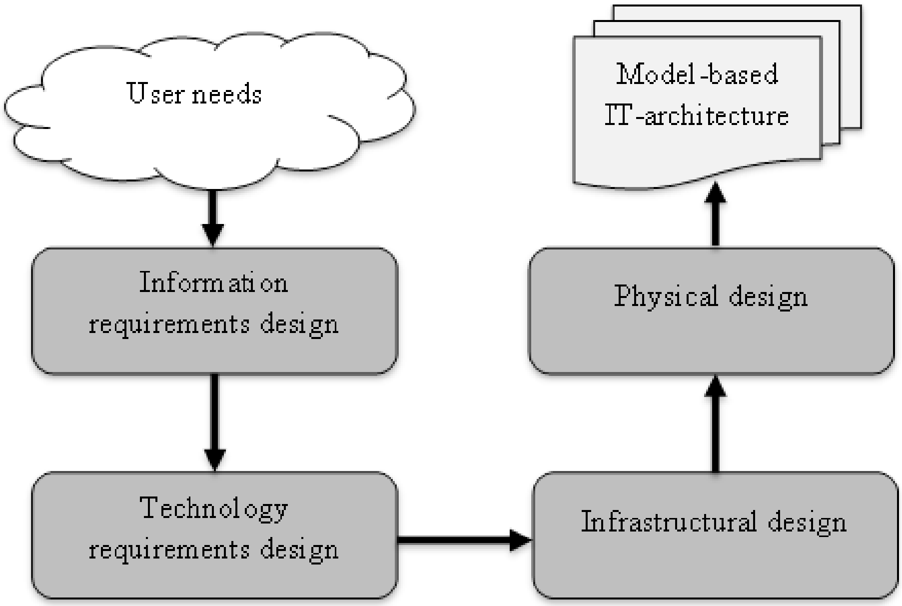


Figure 2.8 *Model for IT architecture design in order to support information systems architecture (adapted from Ardagna et al. 2004, pp. 231-232)*

The methodology by Ardagna et al. (2004) consists of four design phases, each phase producing a sub-model for the IT architecture, i.e. information requirements design phase produces a model for information requirements as an output of the process. Basically, the process start from *user needs*, as depicted in figure 2.8. *Information requirements design* phase creates requirements for organization's information processing capabilities. *Technology requirements design* phase defines the needed IT resources which are needed to fulfill the information requirements that are defined on the previous phase. *Infrastructural design* phase focuses on allocation of the needed IT resources. *Physical design* phase relates to determination of needed hardware, i.e. computers, servers, etc.

After these phases a *model-based IT architecture* is created by four design models (information requirements, technology requirements, infrastructure and physical features). (Ardagna et al. 2004, pp. 231-232) Creation of IT architecture, with or without Ardagna et al.'s model, helps the development of information systems architecture as the IT architecture describes the requirements and the information systems architecture's function is to solve which information systems fulfill the presented requirements and functions, and how.

2.2.4. Data migration

Among information systems and their renewal, data migration process is relevant part of the whole renewal process. Not just information systems and their relations in between, though the data in the information systems must be noticed. Celko (2006, p. 33) presents two so-called truths about organization's data and its environment. First, "no enterprise runs on one and only one database or data source today", and second, "no enterprise database is isolated". For the first truth, Celko (2006, p. 33) states that organizations have different data sources, i.e. desktop databases and various level servers. On the second truth, Celko (2006, p. 33) reminds that organizations have connections to other parties, i.e. customers and suppliers, and those connection are from database to another database or from service to another service.

Thalheim & Wang (2013, p. 1[9]) define the data migration as a process where data from legacy data sources and related information systems is moved to a new system, which is replacing the usage of the legacy system.[10] They also emphasize that the legacy information system and the new information system have different data structures, and therefore the data migration is required (Thalheim & Wang 2013, p. 1). As the data migration is crucially linked to the information systems renewal, it relates to information systems architecture as well, and so, there are some issues that have to be acknowledged when preparing for data migration. Relatively, the data migration requires and benefits from the presence of knowledge management, as it is the data – which is included to the concept of knowledge (see figure 2.1; Sydänmaanlakka 2004) – that is migrated from data sources to new systems. Therefore, it is crucial to manage and acknowledge the important and meaningful data from the organization's business perspective, which needs to be migrated.

Traditionally, according to Thalheim & Wang (2013, p. 2) the data migration consists of three stages: *extraction*, *transformation* and *loading* (ETL). *Extract* includes extraction of the data from the data sources. *Transform* includes the transformation of the old data into a new format and structure – which is required for the new systems – by validating, cleansing and mapping the data. *Load* includes the loading of the migrated data into new source systems. Depiction of the ETL is presented in a figure 2.9. In the figure 2.9, the data from data sources A, B, C and D are extracted for transformation. The transformation section includes data modeling by the given semantics, i.e. validation, cleansing and mapping. After this, the extracted and transformed data from sources A to D are loaded into new data sources, X and Y. (Thalheim & Wang 2013, p. 2)

Apart from ETL and data migration process execution, there are some well-known issues that have to be noticed within data migration. Thalheim's & Wang's (2013) first issue deals with variety of data sources, which can vary in size, format, structure and semantics. And this is result of distinct information systems, which have different data requirements due their nature of functioning. Also it is possible that single information system uses multiple data sources, which complicates the situation. (Thalheim & Wang 2013, p. 1)

---

[9] Thalheim & Wang's article (2013) is in press (corrected proof), so the page numbers are not correct. Though, in case of citation, the page numbers are marked as they appear on the article.
[10] Legacy information system refers to information systems which are widely used though out-of-date and therefore replaced by new information systems.

Based on the previous theory, crucial issue is related to the Thalheim's & Wang's (2013) first issue and is constructed from the theory of knowledge management and information systems architecture (chapters 2.1 and 2.2.3). Before the data migration process begins with extraction of the data, the data has to be mapped and its whereabouts has to be known (knowledge management) and the structure of information systems architecture has to be clear in order to avoid missed systems or data sources (information systems architecture).

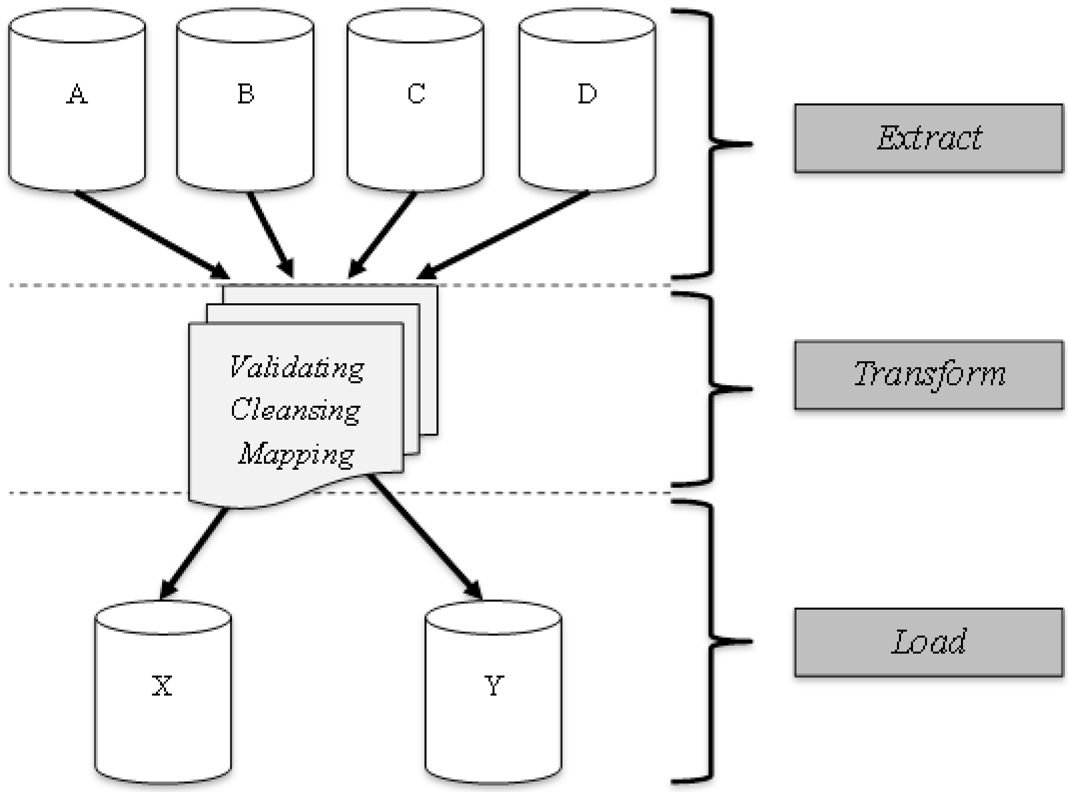


Figure 2.9 *ETL process depiction (adapted from Thalheim & Wang 2013, p. 2)*

Thalheim's & Wang's (2013) second issue relates to migrated data and its quality. The data in information systems and in data sources may be inaccurate, incomplete inconsistent or there may be duplicates of same data presentation. Using the example of ETL in figure 2.9, consider the situation where the data sources A, B, C and D would contain more or less identical data sets or entities. (Talheim & Wang 2013, p. 1)

Though the ETL is done computationally, the transformation may result with unexpected errors as there would be differing presentations of the same data. E.g. after ETL, the data source X could possess two or more presentations of same data, e.g. data for customer's name: *customer_name, customer_first_name, customer_last_name*. Now there would be three data fields for customer's name in data source X, and this is not the desired result for data migration.

Related to the previous example, Celko (2006, p. 34) describes that small- and medium-sized organizations may lack resources – i.e. personnel, time, money, etc. – for data verification and validation, which would be crucial for new data. Therefore the data is usually left untouched in its original format and may lead to errors within data migration, as described above (Celko 2006, p. 34). The described example gives foundation for Talheim & Wang's (2013, p. 2) third issue, in where the data migration should consist iterative execution of data profiling, validating and cleansing.

Through the iterative processing data quality is improved, i.e. no duplicate data, and it also prepares for possible changes in specification. Software and consulting firm Premier International's (2004, p. 2) white paper states that nearly 90 % of original specifications for data migration change, and over 25 % of specifications change more than once. Also, in 62 % of data migration projects, data quality issues were found after implementation of the new system (a study by The Data Warehousing Institute; in Premier International 2004, p. 2). In addition to Thalheim & Wang's (2013) issues for data migration, Kelly & Nelms (2003, p. 506) have presented four key challenges for checking data migration's success:

1. Grabbing the data out of the old and new systems for comparison
2. Differences between the old system and the new system in how the data is stored
3. Changes to live data during the migration process
4. Accuracy of the cut-off parameters that determine whether data is being selected for conversion at all

First, it must be noted that the key challenges by Kelly & Nelms (2003) consider the situation where the data migration has been executed[11] and these four challenges may rise when checking up the quality of the migrated data. The first challenge deals with the issue of how to compare the data between old and new system. As this is a comparison of data – a stage of knowledge – it should be remarked that with application of knowledge management, i.e. proper documentation of migration, the comparison should be straight forwarded and easy. Consider a situation where the data in new system is migrated from several data sources, therefore making the comparison much more complex and time consuming.

The second challenge includes the dilemma of data structure and semantics, i.e. possibly the data fields have different names in the old and in the new system, and some data fields may be excluded from the old system and some data fields may be added to the new system. Surprisingly, the examples of first challenge give illusion that the data in the old and in the new system are identical whereas the second challenges provides more realistic example where the data fields may be excluded/included to the new system. (Kelly & Nelms 2003, pp. 506-509)

[11] Except the key challenge 3, this relates to the time just before and during the data migration.

The third challenge is about minimizing comparison errors between of the data sets in the old and in the new system, while the data migration is executed. The founding idea is to minimize the effect of data changes due to live data, i.e. transactional data which is a result of day-to-day business functioning. For the minimization of live data comparison errors, the data from sources must be downloaded as close to the migration as possible. The fourth challenge deals with data's business relevance for further usage, i.e. how likely the migrated data is not needed, or vice versa, needed data is not migrated. For example, the transactional data has quite short period of time as it comes to the relevance for organization's business making. On the contrary, organization's master data should not be left outside the migration, if the master data is lost for the new system; it affects the functioning of organization's business.[12] (Kelly & Nelms 2003, pp. 506-510)

Generally speaking of migration processes, Breakfield & Burkey (2002, p. 147) have listed typical problems for migrations, and noticeably issue is that the problems do occur within technical and human perspective as well. Though the migration is technical process, the human perspective has to be acknowledged and alike with other projects (alike in knowledge management), the managerial issues should not be forgotten, i.e. metrics and measurement, and quality assurance. (Breakfield & Burkey 2002, pp. 150-151)

[12] For the terms *transactional data* and *master data*, see chapter 3.2 Master data management, for detailed and precise definition.

3. PRODUCT AND MASTER DATA MANAGEMENT

Chapter three focuses on product data management and master data management. These two topics were combined as their close relevance for this research. First, the concept of product data management is discussed in chapter 3.1, which is divided into subcategories within product data management, dealing with the definition, product data systems and challenges, and with the concept of product lifecycle management which relates to the product data management. Second, the concept of master data management is discussed in chapter 3.2, which is divided into subcategories concerning the definition of master data and its management, master data management functions in practice and master data management architecture.

3.1. Product data management

3.1.1. Definition

Saaksvuori & Immonen (2008, p. 7) describe product data as follows: "*physical and/or functional properties of the product – i.e. form, fit and function of the product*". They (2008, p. 7) also note that the product's properties and functions may be described from certain perspectives, e.g. manufacturing, customer, etc. The product data can contain technical data such as blueprints, and it can also contain abstract information such as marketing ideas and slogans. (Saaksvuori & Immonen 2008, p. 7)

The diversity of the term product data management (*later PDM*) is due to that fact that PDM is used as a general term for functions and systems to manage product definition information. Therefore there may be conflicts with terms such as engineering data management (*EDM*), document management, product information management (*PIM*), technical information management (*TIM*) and several others. (CIMdata 1997, p. 2) Due to variety of terms, the term PDM shall be used as a major concept that includes all the functionalities and contents of the ambiguous terms. This was the view of the CIMdata consultancy in the year 1997 but after twelve years, Kropsu-Vehkapera et al. (2009, p. 770) state that the term PDM is still quite new term in the area of academic research and therefore the literature on the discussed topic is scarce.

CIMdata (1997, p. 2) states that the PDM is a tool which is designed to ease the management of product data and its development processes. As the PDM is focused on the management of product data solely, the concept of knowledge management focuses on the distinct stages of knowledge and does not exclude certain data sets, e.g. product data (adapted from Sydänmaanlakka 2004). Therefore, it could be implied that PDM is a special case of knowledge management, where the concept of knowledge management is focused on product data.

Also it is stated by the CIMdata (1997, p. 2) that PDM integrates processes, applications and information that define products in various systems and other applications. Therefore the PDM can be seen as a tool that does not replace current systems, applications or information, but instead it collects and gathers crucial product data in one place. Apart from the PDM's definition, the PDM relates to all product related activities and it is considered as a tool for all whom are working with product data, not solely for managers but for line workers as well (CIMdata 1997, p. 2).

According to Kropsu-Vehkapera et al. (2009, p. 758) the data management is one of the key aspects for company and its business. And the importance of data management becomes more significant as the collaboration between different parties, e.g. partners, suppliers, etc., increase. PDM systems are designed to control and organize the product data in a way that it will enhance company's business processes and supporting functions, which are related to its products. (Kropsu-Vehkapera et al. 2009, pp. 758-759)

As Sydänmaanlakka (2004, p. 186) states that the core idea of knowledge management is to provide sufficient knowledge for the decision making processes (see chapter 2.1.3), the definition of PDM's objective (by Kropsu-Vehkapera et al. 2009, pp. 758-759; in previous reference), is quite similar for its purpose, which on the other hand reinforces the earlier implication of PDM being a special case of knowledge management. PDM is a challenge to companies, despite the fact that do they possess manufacturing functions or not, due to the changing markets and customer needs (Kropsu-Vehkapera et al. 2009, pp. 758-759).

From the perspective of Liu & Xu (2001, p. 251) PDM is related to management and integration of the information that defines the product from design through manufacturing to end-user support. Liu & Xu (2001, p. 251) also state, that the enterprises has become extended and complex nowadays, which has led to constantly evolving and changing IT requirements, which lead to situation where enterprise is divided to independent business units. For the constant development and challenging IT maintenance, Liu & Xu (2001, p. 252) have proposed a product data management to solve this kind of issues, from the perspective of products' and theirs life cycle. As the PDM enables coordination and cooperation between different business units, it can be used globally and it is time and place independent.

With the definitions of CIMdata (1997), Kropsu-Vehkapera et al. (2009) and Liu &Xu (2001), and the additional inclusions of knowledge management by Sydänmaanlakka (2004), the PDM can be described as follows: *PDM is a special case of knowledge* *management, whereas it is a set of processes and functions, which aims to develop, maintain and provide, meaningful and useful product data for organization's product data related business processes and functions, while it encompasses and includes the business related parties and the employees of the organization, in order to enhance organization's business*. (Adapted from CIMdata 1997; Liu & Xu (2001); Sydänmaanlakka (2004); Kropsu-Vehkapera et al. 2009) Though this definition endeavors to include all the issues within the concept of PDM, it should be noted that PDM can be seen differently, depending on the perspective. As the constructed definition views PDM from process-based view, PDM can be seen as a concrete system alike an information system. Therefore, this perspective will be discussed on chapter 3.1.2

3.1.2. PDM system

Liu & Xu (2001, p. 252) highlights that the PDM is strongly related to the engineering and design, but they also note that the PDM is also related to sales & marketing, customer service and project management functions as well. They (2001, p. 252) also make the distinction between PDM and ERP systems, whereas PDM relates to product data through product's life cycle and ERP relates to the manufacturing functions. Therefore the PDM is a precise system for product data and ERP is holistic system for governing and controlling the manufacturing function with the help of PDM system. As the Liu & Xu (2001) define the PDM system as a similar as an information system – which includes and process product data – it should be remarked that the presented theories and concept of information systems in chapter 2.2 can be implied to the concept of PDM system.

The definition of PDM system is not standard but it is seen composed of following components (Kumar & Midha 2006; & Stark 2005, according to Kropsu-Vehkapera et al. 2009, p. 760):

- Information warehouse or data vault
- Information warehouse management: tracing any data related actions
- Document management
- Configuration management
- Product structure management
- Product and workflow structure: definition modules
- Workflow and process management
- System administration management

From the list of PDM systems components, it can be noticed that the linkage between the PDM system and other information systems is crucial for company's business activities and for overall PDM functioning (Wei et al. 2009; & Maletz et al. 2007, according to Kropsu-Vehkapera et al. 2009, p. 760). The research results by Kropsu-Vehkapera et al.'s research (2009, p. 762) showed that the PDM consists of three larger topics: information systems, processes for PDM and product structure. Those three topics correlate with the composed list of PDM components, presented by Kumar & Midha (2006) and Stark (2005) (according to Kropsu-Vehkapera et al. 2009).

Do & Ghae (2011, p. 855) divides PDM systems in five categories based on their purpose of use. Those categories are data vault and document management; workflow and process management; product structure management; classification management; and program management. As Do & Ghae (2011) divided PDM systems in five categories, so did Liu & Xu (2001) too, because the use of same reference for the classification, a report from CIMdata (1997) – *Product Data Management: The Definition – An Introduction to Concepts, Benefits, and Terminology*. The original source states that the classification is done in two broader categories, *user functions* and *utility functions*, and the user functions is divided to the five categories, presented earlier by the Do & Ghae (2011). The adapted presentation by the original source, CIMdata (1997), is presented in figure 3.1.

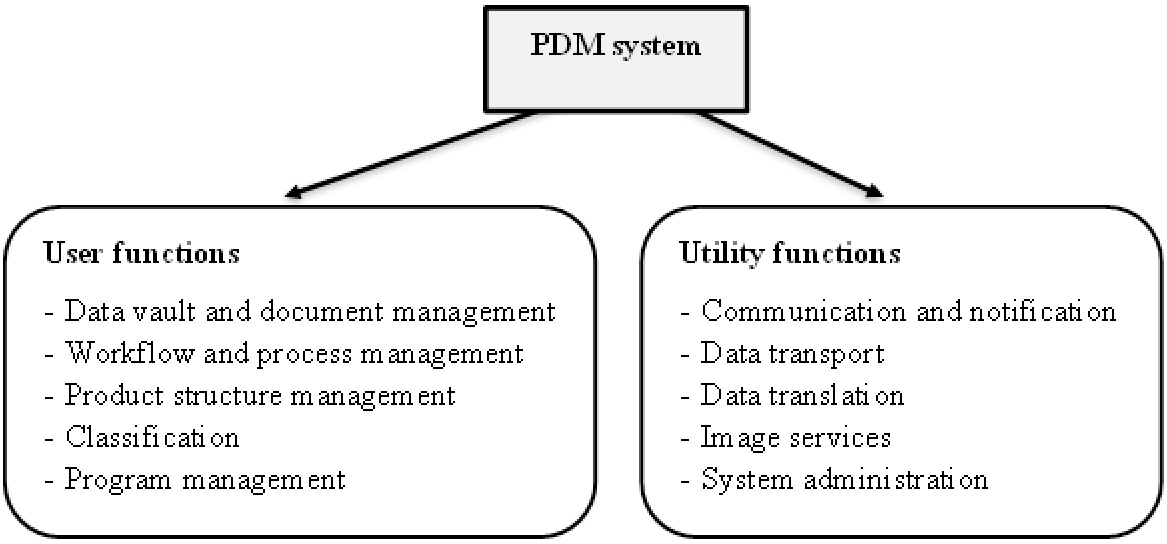


Figure 3.1 *Functionality of a PDM system (adapted from CIMdata 1997, p. 8)*

In figure 3.1 PDM system's functionality is divided into a two distinct categories: user functions and utility functions. User functions are functionalities and capabilities for the end users of the PDM system. Respectively, utility functions are functionalities and capabilities which provide maintenance and support for users and user functions. (CIMdata 1997, p. 8)

Based on the given classification on the figure 3.1, a theoretical model for PDM system should consist of user and utility functions and their functionalities. Data vault and document management provides secure storing and retrieving processes for the data, which is a major factor for correct and up-to-date data. Workflow and process management creates interaction between employees within PDM. This functionality can monitor possible changes to the product data so that the changes will not be unnoticed.

Product structure management relates to the product's structure and the classification enables hierarchical view for the subparts of the product. Program management enables project scheduling and tracking of project goals. (CIMdata 1997, pp. 9-14)

Utility functions are functionalities, which offer usability to the end-user as in notifying all new changes and events, providing the required data, enabling possible translations, handling the data as unified format, authorization, data back-ups, user interface layout, new functionalities and so on. (CIMdata 1997, p. 8) By these functionalities and requirements for PDM system, a theoretical model for PDM is presented in figure 3.2.

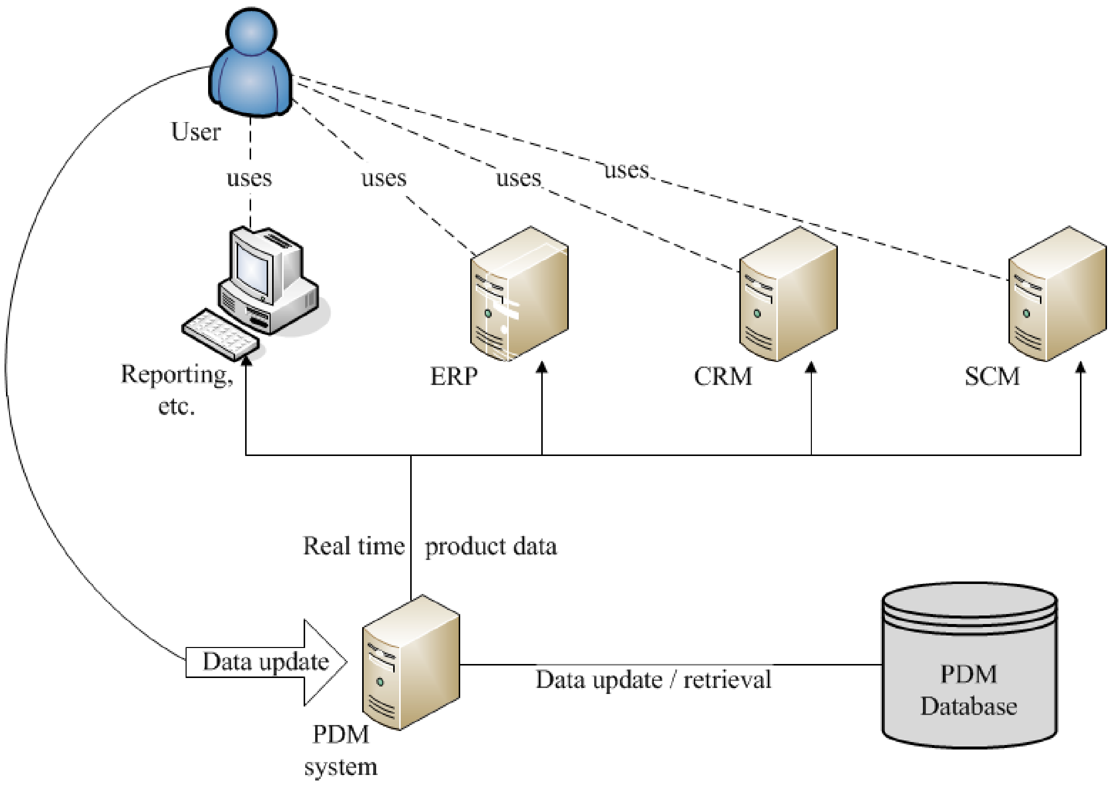


Figure 3.2 *Theoretical model for PDM (adapted from CIMdata 1997 & Kropsu-Vehkapera et al. 2009)*

In figure 3.2 there is presentation of a theoretical PDM-model, which is an ideal solution and it is not a real in any context. The PDM-model depicts ideal situation where user uses different applications and systems, such as reporting, ERP, etc. For any updates concerning the product data, the updates are done in one place, which is the PDM-system. By update action is meant actions such as insertion, deletion, correction, etc.

Continued from figure 3.2, due to the update action in PDM-system, the updated data will be stored in PDM-database, where all the relevant product data is stored and maintained for further use. The PDM-system also has another functionality to provide up-to-date product data for applications and systems using the product data. By this kind of arrangement, the product data updates are centralized to one point in the system and it reduces multiple update processes in various applications and systems.

As mentioned previously, the PDM-model is an ideal situation and the depiction of the model in figure 3.2 does not take a stand on the execution and implementation of the depicted model. For example Liu & Xu (2001, p. 255) propose the idea of harnessing the power of Internet and making the PDM-environment web-based. And this idea was presented back in the year 2001 and nowadays this solution is quite standard in all levels. For theoretical examination of the execution and implementation of PDM system, e.g. the waterfall model by Haikala & Märijärvi (2006) can be used in together with the information systems architectural designs (see chapter 2.2.3). Apart from the technological perspective, Eynard et al. (2006, p. 341) proposed the use of unified modeling language (*UML*) for modeling, specification and implementation of PDM-system, especially within product structure and workflow activities.

3.1.3. PDM challenges

As the previous chapters described the definition for PDM and PDM-systems, those descriptions were more likely positive and ideal representations of the PDM and its content. On contrary, in this chapter the possible challenges and obstacles of PDM are described and discussed to have more profound insight about the world of PDM. The research conducted by Kropsu-Vehkapera et al. (2009) dealt with four large high-tech companies and their practices and challenges with PDM. Though, the size of the examined population is quite small for overall generalization for consistent theory building, it still gives sufficient details about PDM challenges that may be encountered. The list of PDM challenges are as follows (Kropsu-Vehkapera et al. 2009, p. 764):

- Product data availability, without single PDM application
- Data quality: has to be up-to-date, reliable and relevant
- Users do not produce or maintain product data
- Combining different applications for the single PDM system
- Lack of standardized processes related to product data
- Lack of solid definition for PDM in organization
- Personnel changes
- Changes in product structure
- Standardized product data

Due to findings of the research conducted by Kropsu-Vehkapera et al. (2009), it can be seen that the challenges with PDM systems are not simply technological issues[13]. For example, there exist issues related to the personnel and organizational features as well. In table 2 is the rearrangement of the PDM challenges list, which classifies each challenge to relevant categories.

Table 2. *PDM challenges by categories*

	Category		
Challenge	*Organizational*	*Personnel*	*Technological*
Product data availability	x	x	x
Data quality		x	
Lack of user involvement	x	x	
Efforts for single PDM system	x		x
Lack of standardized processes	x		
Lack of solid definition for PDM	x		
Personnel changes		x	
Changes in product structure	x		x
Standardized product data	x	x	x

In table 2, is a categorization of the PDM challenges by three factors: organizational, personnel and technological. Organizational factor relate to issues within organizations functions and ways of work. Personnel factor relates to the employees of the organization and the technological factor relates technology issues as applications and their restrictions. The division of the PDM challenges into a three categories is not directly based into the work of Kropsu-Vehkapera et al. (2009), thus the categorization is done by the knowledge of the author himself with applying Kropsu-Vehkapera et al.'s (2009) research results.

As mentioned before (in chapter 3.1.2), Kropsu-Vehkapera et al.'s (2009) research findings defined that PDM consists of three larger entities; information systems, processes for PDM and product structure. Respectively, the *information systems* relate to technological category, and, *processes for PDM* and *product structure* relate to the organization and personnel, which was divided into two separate categories. Due to this categorization, the core reasons for this kind of an arrangement are briefly summarized in the table 3.

---

[13] As the research focus emphasizes the IT-perspective (see chapter 1.3), it was expected, that non-technical and non-IT related issues may emerge within research. Therefore they are taken into consideration as it serves the objective of the research and enables enhanced understanding for the research problem and for the given recommendations.

Table 3. *Explanation of the challenges and their relation to different categories*

Challenge	**The relation of the challenge and the category**
Product data availability	Product data availability depends of all three factors as a shortage in one factor leads into overall inferior availability.
Data quality	Data quality is mainly personnel factor as it depends upon the employees and their understanding for quality data.
Lack of user involvement	Organization and personnel have to work together for greater good and motivate the users.
Efforts for single PDM system	Depends on organizations investment willingness and is it even technologically possible or plausible on the other hand.
Lack of standardized processes	Organization has to develop standardized processes for the employees in order to achieve functional PDM.
Lack of solid definition for PDM	The employees have to be sure of the true meaning and impact of PDM so that they can work according to organizations definition and ideals.
Personnel changes	Due to several reason personnel will change no matter what, but instead this situation should be foreseen until it is too late.
Changes in product structure	Organization will have to change or merely adapt its offering for changing customer needs and therefore product structures may change and affect technological areas as well.
Standardized product data	The product data in organizational, personnel and in technological level has to be standardized in order to ease to communication and business functions.

In table 3 is definition of the categorization factor and the PDM challenge and their relation between. To sum up the findings of Kropsu-Vehkapera et al. (2009) and the tables above, the PDM challenges may be encountered from different perspectives, e.g. personnel and organization besides the obvious technological perspective. Due to these different perspectives and the earlier implication of PDM being a special case of knowledge management, plain application of knowledge management is not enough, as the technological perspective requires application of information systems science, such as development models and architectural designs as well (see chapter 2.1 & 2.2). Also, it must be noticed that the impact on one presented factor may affect the other factors as well. Therefore, one could say that the holistic entity is larger than the sum of its parts.

3.1.4. Product lifecycle management

Eynard et al. (2006, p. 331) emphasize that in addition to product data and its management and storing, the information related to its lifecycle has to be acknowledged also. Alike Eynard et al. (2006), Gimenez et al. (2008, p. 213) describe correspondingly, that the concept of product lifecycle management (*later PLM*) is among other information gathering and processing systems. The CIMdata (2002, p. 2) mentions that the concept of PDM was used firstly when referred to product data but the concept evolved to PLM, which has its foundation in ideology of PDM. Also Saaksvuori & Immonen (2008, p. v; *see Preface*) mentions that the concept of PDM is narrower than the current concept of PLM. It seems inevitable that the concept of PDM overlaps with the concept of PLM, so similarities and differences of these concepts are discussed in this chapter.

Cantamessa et al. (2012, p. 193) describe that the concept of product lifecycle management (*PLM*) was an extension to the previous concept of product data management (*PDM*), which has been discussed earlier. The PLM is seen as a connective factor between product-related knowledge across the organization and its interest groups. (Cantamessa et al. 2012, p. 193) Saaksvuori & Immonen (2008, p. 9) describe the PLM as a large entity consisting of functional methods and processes in order to manage and control product information through product's lifecycle. They (2008, p. 9) also point out the fact that the PLM is not a single application or method but a holistic entity as described above. Though, they (2008, p. 13) also mention that the term PLM is usually used to describe a PLM system, not the methodology and the ideals behind the concept. (Saaksvuori & Immonen 2008, pp. 9; 13)

As Cantamessa et al. (2012) and Saaksvuori & Immonen (2008) defined the concept of PLM, so did Ameri & Dutta (2005, p. 578), which also see the PLM as a concept that includes creation and dissemination of product related information throughout the organization. Also Ameri & Dutta (2005, p. 578) state that the inability of PDM systems to function with ERP/CRM/SCM (*Enterprise Resource Planning/Customer Relationship Management/Supply Chain Management*) systems led to the development of PLM and PLM systems which could function beyond PDM systems limitations. They (2005, p. 578) describe that the PDM systems focused only on managing product data, as in engineering data such as CAD (*Computer Aided Design*) drawings, whereas PLM systems relate to knowledge management and the overall business functions as in sales and marketing.

CIMdata (2002, pp. 2-3) states that the concept of PLM overlaps with the concept of PDM and they can carry multiple meanings. It is the case as there is no apparent distinction between these concepts. The definitions presented above are very clear, though they may have different meanings to different interpreters. Despite the overlapping, CIMdata (2002, p. 3) depicts the PLM in similar way as the other authors above. Also, it is noted that PLM and the PLM systems are far more extensive than PDM, which on the other hand has led to the dominance of PLM instead of PDM (CIMdata 2002, p. 3).

Ameri & Dutta (2005, p. 586) present that the PLM is more commonly thought as a collection of IT tools, which on the other hand complicates PLM's knowledge management perspective. This is a similar statement as the Saaksvuori & Immonen (2008) presented about PLM mistakenly thought as PLM system, not a concept. For the organization's perspective, it should be clear what they mean with different concepts and terms, so the confusions can be avoided. At the end, if organization's processes and systems are not working as planned, it does not matter what terminology or concepts are used.

3.2. Master data management

3.2.1. Definition

Loshin (2009, p. 6) defines master data as business objects, consisting of data themes and information, which are widely used across organization in different applications and the master data objects (e.g. customer object as in customer and the data related to it) are those that matter the most. The master data can be about customers, products, suppliers, vendors, etc. and the hierarchy of master data may vary depending on the organization and its business, e.g. product may consist of other products. Master data is usually related to several business functions across the organization and it tends to be static in nature, i.e. the master data changes rarely, e.g. customer's or product's name. (Loshin 2009, pp. 6-8)

The need for master data and its management derives from complexity of organization's structures and applications. With this statement, Loshin (2009, p. 2) refers to the situation where organization's applications are referring to various sets of data, which are supposed to represent the same information. For example, the instance of customer in sales and in finance applications may vary in detail and in content. For clarification, figure 3.3 presents the idea of similar but yet distinctive data sets for customer data. (Loshin 2009, p. 2)

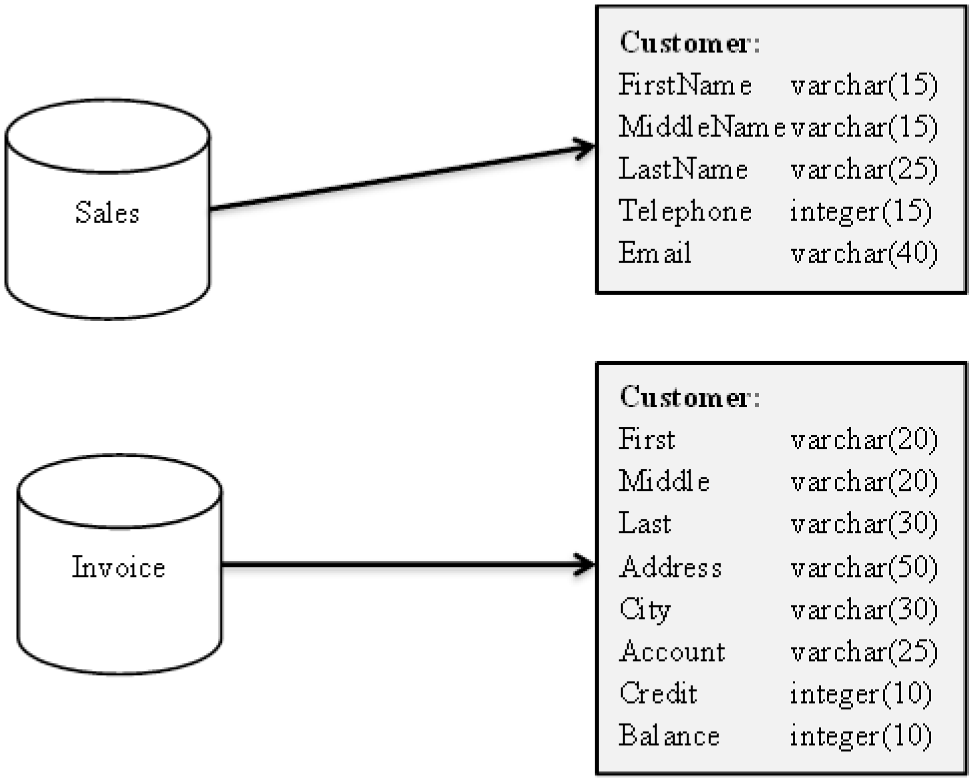


Figure 3.3 *Customer data set differences and similarities between different business applications (adapted from Loshin 2009, pp. 2-5).*

Obviously, different business functions have different needs for their business activity. In figure 3.3 is presented *Sales* and *Invoicing* functions and their customer data sets (*later data object, a term used by Loshin 2009*). As the figure 3.3 shows, both functions need information about the customer, but the customer data objects have distinctive data contents and definitions. The sales function requires only customer's name and contact information i.e. telephone number and email address. In addition to the customer name, the invoice function requires invoicing address and account numbers, as well the current credit and balance information.

Also, in figure 3.3, the functions have different definitions for the customer data, e.g. the name of the customer varies in length of the name. The sales define the first name to be maximum of 15 characters and the invoice defines it as 20 characters long ("varchar" refers to variable character field, "integer" refers to numeric data and the number in parentheses refers to the maximum length of the data field, this kind of notation is used e.g. in database specifications). This is the situation as Loshin (2009) described where both functions refer to the same customer, though the information contents vary in some detail. From the perspective of data migration (chapter 2.2.4, e.g. Celko 2006; Thalheim & Wang 2013), variations in data presentation may lead to unwanted errors in data migration, where the data is compared computationally in order to avoid duplicate data.

This contradiction of disparity of customer data objects may lead to problems within the business functions as the data is redundant and it raises the question of reliability: which one of the data object presentations is up to date and are there any defects? This is the problem from user's perspective, but there exists problem with the maintenance naturally. With two or even more separate data object presentations, there has to be as much of update actions to keep the data up to date, e.g. if there is seven distinct presentations of customer data, then in case of data update, the update has to be done seven times. This sounds terrible situation and Loshin (2009, p. 2) describes this as the challenge of master data management (*later MDM*).

The first issue with MDM is to identify and define those critical business objects and data sets which are important for organization's business and functioning. The next step is to integrate those data objects to a uniform format which will be the basis for master data objects. After this, the master data objects has to be made available for organization wide usage, e.g. for applications. (Loshin 2009, p. 2) So it could be stated, that the objective of MDM is to create, maintain and distribute the crucial master data objects for organization's use. Though it seems quite simple and straightforward process, it still needs management and governing, as it is not static but a dynamic process which changes and differentiates through organizational and business alterations. (Loshin 2009, p. 2) As the figure 3.3 presented the starting point for MDM adoption, the result of data integration for situation in figure 3.3 is depicted in figure 3.4.

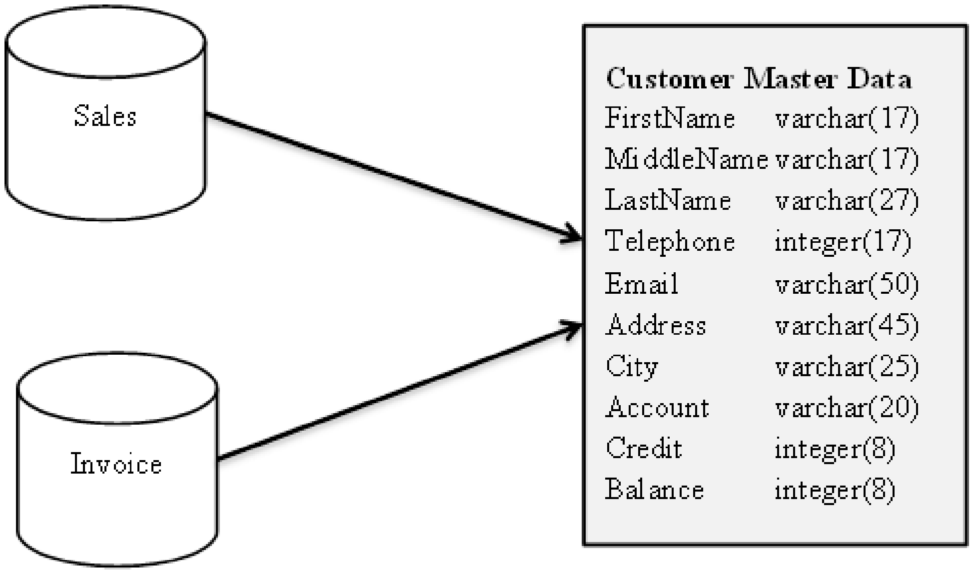


Figure 3.4 *The unified view of customer data, i.e. master data for different business functions (adapted example from figure 3.3, for further detail see Loshin 2009 or Loshin 2011).*

In figure 3.4 is a situation where customer data objects are integrated to a single customer master data object. The customer master data is exploited by sales and invoice function together, instead of having two different customer data objects for each business function.

The figure 3.4 is quite simplified presentation of the solution for the customer master data as there may be more data fields than presented and there may be more business functions using this specific master data object. Based on these examples, figures and Loshin (2009), it can be stated that the MDM: "*is a set of processes and actions that aim to create and maintain unified view of business critical data objects, which enable effective business for the organization*".

Relatively to the objective and definition of PDM and knowledge management, the MDM has a few features in common. First, PDM focuses only on the product data and its management (see chapter 3.1, e.g. Kropsu-Vehkapera et al. 2009), whereas the MDM focuses on the master data, no specific topic, and its management. Though, if the MDM focuses on the product master data, then the PDM and MDM could be almost identical. Thus, the most obvious difference is that the PDM includes all product related data and MDM includes only business critical data. However, in a special case, organization's product data can be business critical, so it cannot be excluded, and therefore PDM and MDM would have same data sets to cover for their objectives. Second, MDM includes the business critical data; without exclusion of data topics – e.g. products, customers, etc. – so alike the PDM, MDM can be implied to be a special case of knowledge management. Though, if the MDM includes business critical data, then does the knowledge management include the non-master data as well?

3.2.2. MDM in practice

Loshin (2011, p. 330) emphasizes that the MDM is a collection of data practices which include managerial and technical perspectives. Alike PDM – described in the chapter 3.1 – the MDM has to deal with issues of governance and information technology combined, not just the other perspective. From a practical perspective, Loshin (2009, p. 9) defines that the MDM "*incorporates business applications, information management methods, and data management tools to implement the policies, procedures, and infrastructures that support the capture, integration, and subsequent shared use of accurate, timely, consistent, and complete master data*". Based on to the Loshin's (2009) holistic and comprehensive definition, the described MDM process is depicted in figure 3.5.

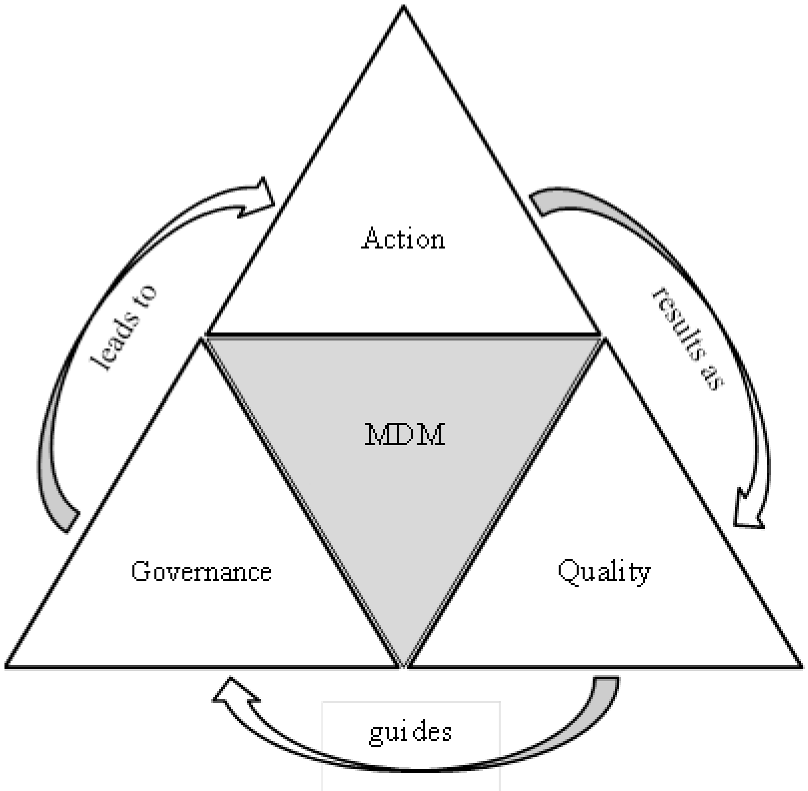


Figure 3.5 *The MDM process (adapted from Loshin 2009, p. 9)*

The MDM process is depicted in figure 3.5 and it consists of three distinct functions: governance, action and quality. First, governance sets the framework and guidelines for MDM and therefore leads to action. Second, action relates to the everyday work and functioning for improved MDM functioning and it can be measured within quality of MDM. Third, quality compares the given and the accomplished milestones, and gives feedback and guidance for MDM governance, thus process begins again and keeps on going. (Loshin 2009, p. 9) For clarification, the figure 3.5 depicts a process-based view of MDM and those three functions are performed more or less parallel than separately and in a series.

Parting from Loshin's (2009) MDM process, Otto (2012) studied Bosch (the company) and presented their MDM reference framework. Bosch's framework consisted of four components: *Governance system*, *Master data provisioning*, *Master data utilization*, and *Master data concepts & projects*. Also the model includes the division responsibilities to organizational/functional responsibility and technical responsibility. The framework is depicted in figure 3.6.

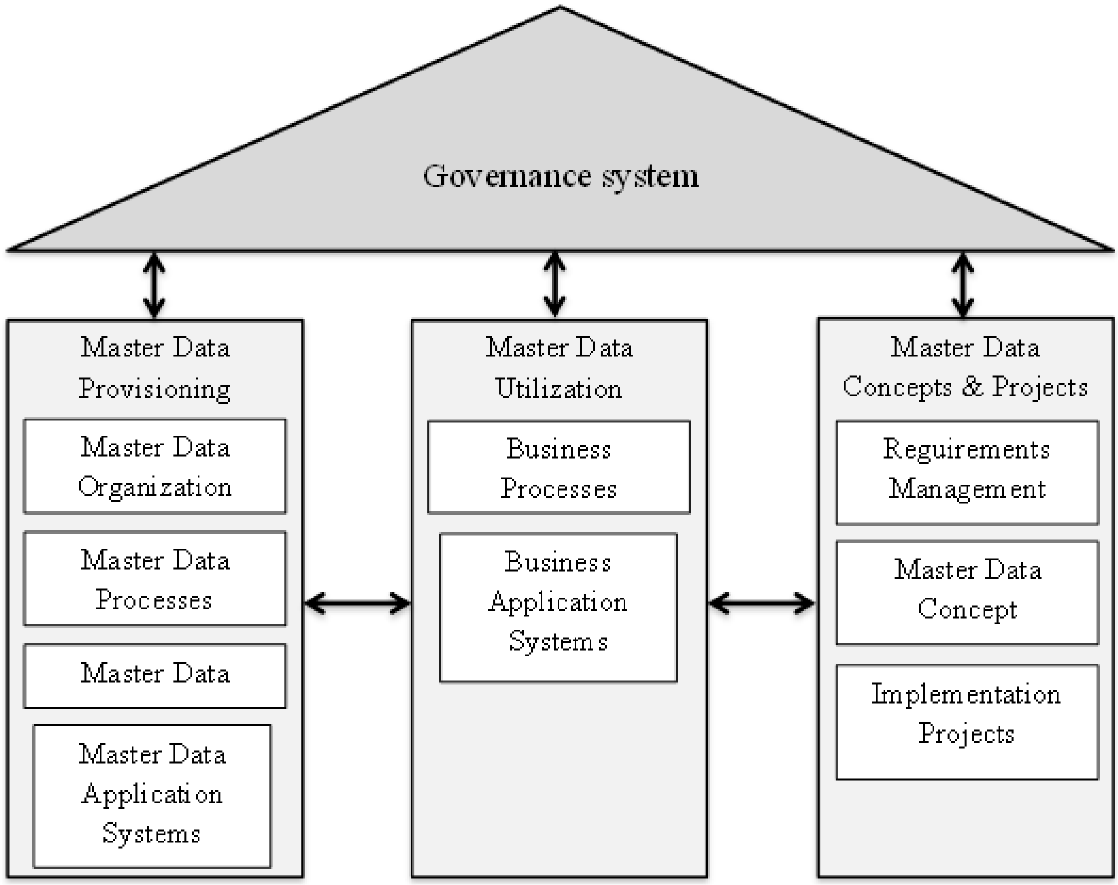


***Figure 3.6** MDM reference framework at Bosch (adapted from Otto 2012, pp. 339-340)*

At the top of Bosch's MDM framework in figure 3.6 is the *governance system*, which contains master data strategy and master data controlling for master data objects and their quality. Under the governance system are the provisioning, utilization and concepts & projects functions for master data. *Master data provisioning* includes design and maintenance of master data such as processes for creation or update, design for the master data model and the applications for master data storing and distribution. *Master data utilization* deals with the usage of master data in business processes and the use of business applications which support business processes. *Master data concepts & projects* relate to the requirements and specifications of master data and their implementations on organizational and technical level, usually focus is on the lifecycle of the master data objects. (Otto 2012, p. 340)

Derived from Bosch's MDM framework, it should be noted that the MDM requires vertical and horizontal cooperation within the organization. That is, the governance functions vertically cooperating with different business departments implementing the MDM framework and the business departments cooperate together horizontally. (Otto 2012, pp. 339-340) Loshin's (2009) MDM process, in figure 3.5, seems more periodical and simple, than the Bosch's framework, though the functionality of the MDM process is more important than its appearance. Loshin's (2009) model provides upper-level framework whereas Bosch's framework has more detail and distribution of liability.

As the basis of the MDM is to create and maintain unified view of business critical data objects, as mentioned before, therefore it is crucial to recognize these objects and keep inventory of them. Also the master data objects require definition of semantics and hierarchies, and relations between different objects. On top of these actions, a framework is needed for management and governance. Loshin (2009) emphasizes that the greatest obstacle for MDM success are not technical issues, rather the organizational issues are the ones hindering the MDM success. Moreover, these organizational issues are related to the change process (i.e. implementation of MDM) and its management, and they can be avoided by proper preparing and instructing the employees for upcoming events. (Loshin 2009, pp. 14, 16-17)

As the practical process-view of MDM is presented, then one may ask that what are the true benefits of MDM and is it worth of trying. For the benefits of MDM, Loshin (2009, pp. 10-12) has listed up to thirteen different benefits of MDM:

- Comprehensive customer knowledge
- Improved customer service
- Consistent reporting
- Improved competitiveness
- Improved risk management
- Improved operational efficiency and reduced costs
- Improved decision making
- Better spend analysis and planning
- Regulatory compliance
- Increased information quality
- Quicker results
- Improved business productivity
- Simplified application development

According to Loshin (2009, p. 10) the MDM will be greatly beneficial for organization, but it should be recognized that the master data integration and management are not the end objectives, instead they stand to support other business activities and functions to succeed in a intended fashion, alike PDM. Also it should be clear that the benefits of these thirteen factors should be measured and monitored to ensure the gained advantage.

3.2.3. MDM architecture

According to Loshin (2011, p. 339) the possible MDM architecture depends on the business requirements and the usage purposes for master data. Loshin (2011, pp. 336-338) divides the master data usage purposes into a three scenarios: *reference information management*, *operational usage*, and *analytic usage*. *Reference information management* relates to importation of data into a master data system, where the data is refined for further usage, publication and distribution. *Operational usage* refers to usage of the master data in the master data system and the *analytic usage* refers to usage of master data for decision support and business intelligence operations. From the perspective of these master data usage scenarios, three MDM architecture models are presented (Loshin 2011, pp. 336-339):

- Registry (see figure 3.7)
- Full repository (see figure 3.8)
- Hybrid approach (see figure 3.9)

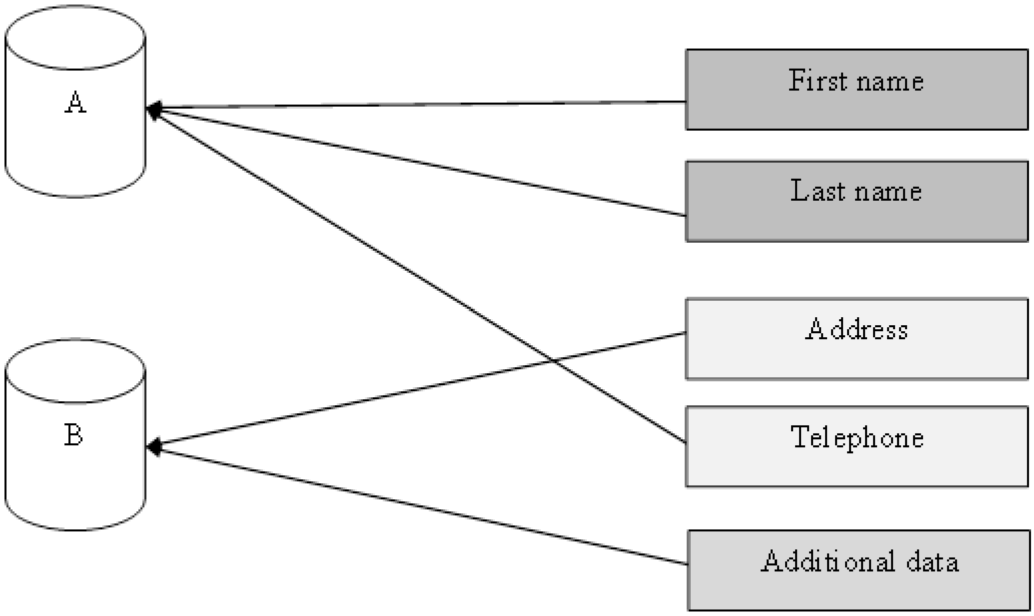


Figure 3.7 *MDM architecture: registry (adapted from Loshin 2011, p. 341)*

In figure 3.7 is a presentation of *registry* type MDM architecture. In this figure the registry is on the right side where the master data mapped to the existing data assets, e.g. databases A and B. All the registry data have a unique index number which is used to identify the data content. By looking the figure 3.7, it can be observed that the name entity, consisting of first and last name, is found from the database A. With this information of master data whereabouts, the relevant business applications can exploit the desired information from database A by using the registry and its indices. (Loshin 2011, pp. 340-341)

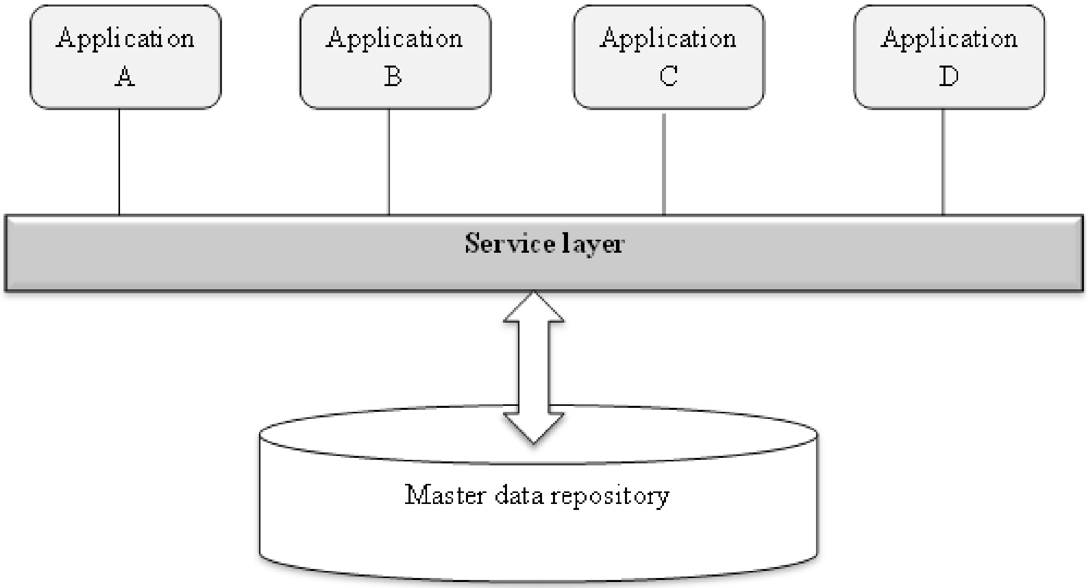


Figure 3.8 *MDM architecture: full repository (adapted from Loshin 2011, p. 342)*

In figure 3.8 is a presentation of *full repository* type MDM architecture. In full repository style architecture, all the data is stored in single master data repository as seen in figure 3.8 and the service layer acts as an access point to the master data. E.g. application A requires master data for business functionalities: the application A sends a request to the service layer and then retrieves the required data from repository and transmits the data to the application A. It is crucial to notice that there are no copies or replicas of the master data within the applications. Therefore the service layer serves a vital function for business function and operations. (Loshin 2011, pp. 341-342)

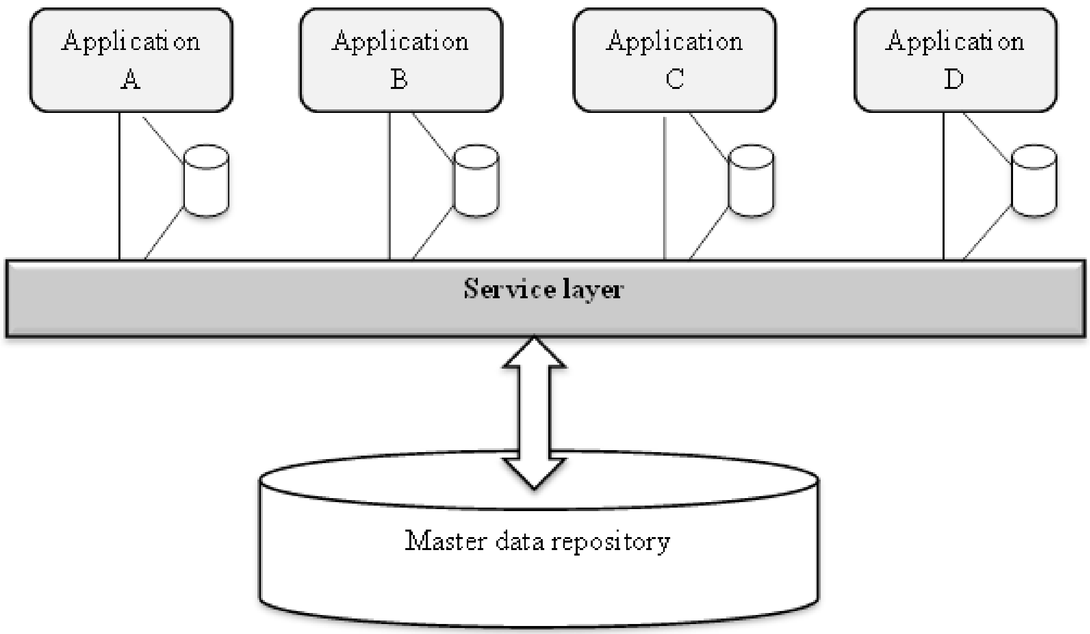


Figure 3.9 *MDM architecture: hybrid approach (adapted from Loshin 2011, p. 343)*

In figure 3.9, is a presentation of *hybrid approach* type MDM architecture. This architecture type is called hybrid approach as it tries to be a trade-off between registry and full repository architectures. Distinction between registry and full repository is that in the figure 3.9 there are databases for every application in addition to the master data repository. The idea of hybrid approach is to distribute master data objects to the applications periodically and the applications exploit their own databases. Also, on a timely basis, the application based master data is extracted to the master data repository and is both consolidated and cleansed for further use. On the other hand, another approach may be taken whereas the service layer has more important role when it manages the applications need and usage of master data. I.e. application based data is exploited from its own database whereas the common or public master data objects are retrieved from master data repository. (Loshin 2011, pp. 342-343)

Otto's (2012) case study at Bosch reveals four distinct MDM architecture types which are: *analytical approach*, *transactional approach*, *coexistence* and *parallel approach* (figures presented by Otto 2012, see figure 3.10). These types have similar characteristics as Loshin's (2009) three architecture types. Otto (2012) describes the architectures with source systems (applications providing master data), target systems (applications exploiting master data) and master data server, which is same as Loshin's master data repository (*later repository*).

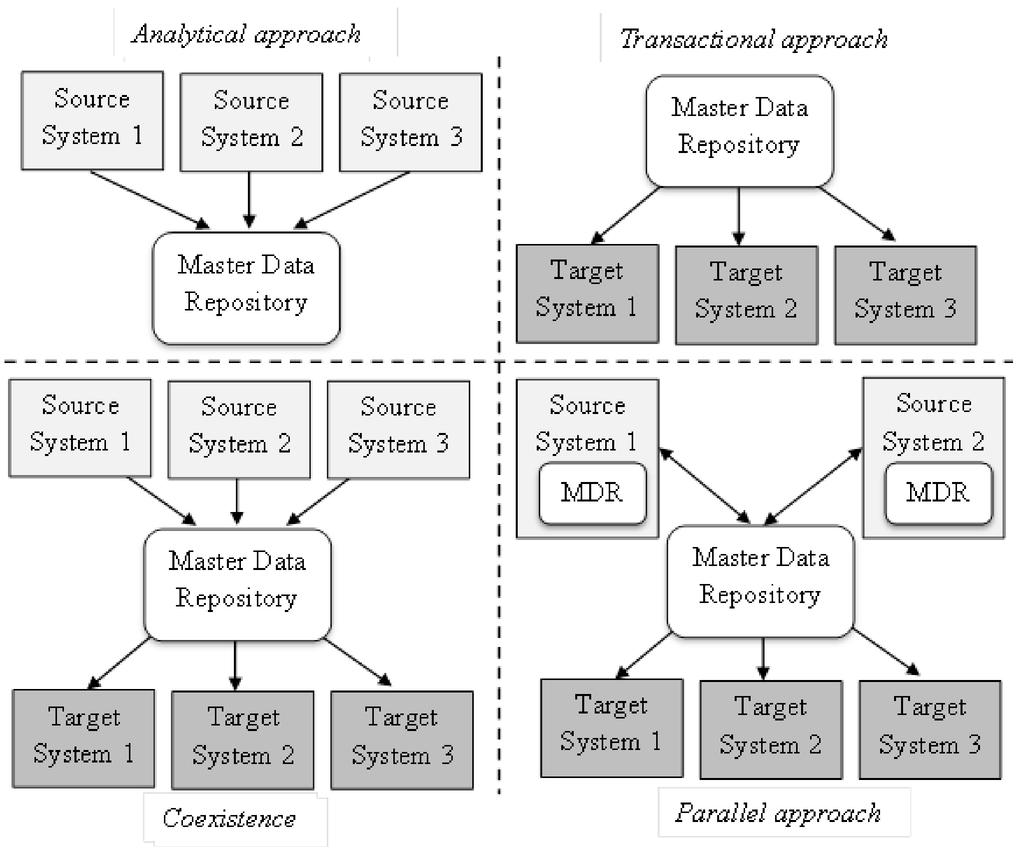


***Figure 3.10** Four MDM architecture types from case study at Bosch (adapted from Otto 2012, p. 342)*

In figure 3.10, are the four MDM architectures presented by Otto (2012). In the left upper corner is the *analytical approach*, in the right upper corner is the *transactional approach*, in the left lower corner is the *coexistence* and in the right lower corner is the *parallel approach*. As said before, the MDM architecture types by Loshin (2009) and Otto (2012) have similarities, such as the *full repository* and *transactional approach* are alike. The *hybrid approach* is a combination of *coexistence* and *parallel approach*. Loshin’s (2009) *registry* differs from other architectures as it seems to be more like a catalog for master data and its whereabouts.

Continued with figure 3.10, *analytical approach* is an architecture where source systems i.e. local applications create and maintain master data and the master data is periodically transferred to master data repository. Also the analytical approach is used when preparing for data migration projects. *Transactional approach* is similar to Loshin’s (2009) *full repository* –architecture, where master data repository provides master data for target systems. *Coexistence* is a mixture of analytical and transactional architecture types, where the source systems provide the master data to the repository and the repository distributes the master data for target systems. *Parallel approach* is quite similar to the coexistence type, although there some detailed differences e.g. in the master data consolidation, done by local users or experts. This may sound vague in written, yet there is another difference between the coexistence and parallel approach; within parallel approach the master data is transmitted back and forth between source system and the repository, whereas coexistence type allows master data to go from source system to the repository, not vice versa. (Otto 2012, pp. 341-342)

4. CASE COMPANY, RESEARCH METHODS AND PERFORMANCE OF THE RESEARCH

4.1. Case company

4.1.1. Imperial Tobacco Finland Oy

Imperial Tobacco Finland Oy (*later ITF*) is subsidiary of Imperial Tobacco Group PLC (*later ITG*). ITF is responsible for importation, warehousing, distribution and marketing of the goods sold in the Finnish market. ITF has offices in Kaarina, Helsinki and also sales representatives around Finland. ITF's revenue for fiscal year 2011 was 39 million (€) and ITF employs 39 person.

In 2011, ITF was the second largest player in the Finnish market with its share of 26,6 %. The markets consist of four segments: cigarettes, roll-your-own (*later RYO*), cigars and supplies. In 2011, market share per segment was for cigarettes 20,4 %, RYO 64,0 % and cigars 43,3 %.

4.1.2. Imperial Tobacco Group PLC

Imperial Tobacco Company was founded in 1901 consisting of several tobacco companies which encompassed geographical areas of Great Britain and Ireland. Imperial Tobacco Company was founded as a subsidiary of American Tobacco Company, because of the unsuccessful attempts to buy small tobacco companies one by one. Through various phases and ownership changes, in 1996 Imperial Tobacco Company was separated from its parent company Hanson PLC and continued its business as an individual company, Imperial Tobacco Group PLC. From that point on, the ITG continued its business and expanded its operations globally.

In 2002, ITG acquired Reemtsma Cigarettenfabriken GmbH and it was strategically significant milestone for ITG on its way to become a global tobacco company. This acquisition strengthened ITG's position in Germany, Central and Eastern Europe, and Asia as in supplement for ITG's former positions in UK, Western Europe, Australia and Africa. Until 2008, ITG purchased the world's fifth largest cigarettes manufacturing company Altadis through corporate acquisition. Since then, the Finland's office Altadis Finland Oy was known as Imperial Tobacco Finland Oy. In 2011, ITG was world's fourth largest company in the tobacco industry and is the global market leader in cigars and accessories as well as in RYO and pipe tobacco. ITG operates in 160 countries and has personnel over 38000 employees.

4.1.3. History

Imperial Tobacco Finland Oy’s history begins from 1845, when tobacco company P.C. Rettig was founded in Turku, Finland. In 1995 P.C. Rettig sold its business to R.J. Reynolds International and acquisition included cigarette and RYO factory in Turku and cigar factory in Pietarsaari.

In 1998 R. J. Reynolds manufacturing functions were closed in Finland, sales and marketing organizations were sold to French Seita S.A. and Finnish office was named Seita Tupakka. Cigar factory in Pietarsaari was sold to Swedish company Match. In 1999 Seita Tupakka moved from Turku to Piispanristi, Kaarina. In 2000, after merger of Seita S.A. and Tabacalera, Seita Tupakka continued its business under the name of Altadis Finland Oy. After a corporate acquisition in 2008, Altadis Finland Oy was merged to ITG and Finland’s office was changed to Imperial Tobacco Finland Oy.

4.2. Research methods

4.2.1. Data collection

As it was mentioned in the first chapter, the research methods refer to data collection and data analysis techniques. Ghauri & Grønhaug (2005, p. 108) describe that the data collection technique will depend on the need of the data, i.e. what kind of data is needed for solving the research problem. In figure 1.1, on page 9, were different research methods as interviews, surveys and observations. In this research it is crucial to understand, explore and make explanations of the ITF’s functioning. As the research problem is quite structured and requires detailed information e.g. product data content, the choice of observation is not suitable for this research. If the research problem would be about working environment and an efficiency of certain physical procedures, then the observations could deliver more insight than other methods.

With the methods of interviews and survey, the interviews seem to be more appropriate method than survey as the survey method is more likely linked to survey strategy as a strategic choice (Saunders et al. 2009, p. 360). Also the survey is a great technique for solving customer satisfaction though the survey does not tell us why the customers are satisfied. Instead, interviews can be used to gain in-depth insight about studied phenomena and understand its functioning. (Saunders et al. 2009, p. 362). Noteworthy issue is the size of the case company, approximately 40 employees, that is more suitable for interviews than survey, which would be more effective with larger population. This is a significant benefit for choosing the interviews as a data collection method. Interviews give more profound insight about company’s structure and processes and it is more suitable method for clarifying the vague answers.

Saunders et al. (2009, p. 320) divide interviews between structured, semi-structured and unstructured interviews. Structured interviews are interviews consisting of identical questions, despite the interviewee and this type interviews are usually related to quantitative research. Semi-structured interviews consist of predetermined questions, as in structured interview, and themes to be asked about. The questions can vary between interviewees but the interview deals with same themes and same basic questions. Unstructured interviews do not rely on any predetermined questions or themes. It is more like a free conversation, though the questions and topics have to be related to the reason of the interview, i.e. the research problem. (Saunders et al. 2009, pp. 320-321)

Each interviewing type, presented by Saunders et al. (2009), has strengths and weaknesses, which has to be considered when conducting a research. Structured interview requires careful planning of the questions so the questions would yield usable results. When planned poorly, the results tend to be poor and they lack of usability. On the bright side, the advantage of structured interview is the analysis of the data, which can be done with quantified methods.

On the contrary, unstructured interview can reach the core reasons for understanding the problem and it can deliver insights that are impossible with structured interview. Downside of unstructured interview is the freedom of choices: when the interviewer has scarce experience about interviewing, the interviewer may not stick to the point and as a result the findings may be shallow. The semi-structured interview is a trade-off between these two interviewing types and it gives the interviewer a core questions for the interview and the freedom to apply his/hers knowledge to the interviewing situation. (Saunders et al. 2009, pp. 320-321)

Although there are only three different interviewing types to choose from, it is not straightforward to choose one over others. As described above, each interviewing type has advantages and disadvantages, so the presence of researcher’s experience and the research problem guides the selection of interviewing type. With researcher’s low experience the choice would be the structured interview though constructing the questionnaire requires experience and know-how. Therefore in this research is used combination of structured and semi-structured interviews. By this choice there are certain themes used for interviews and there exist predetermined questions, which form the foundation for the questionnaires. The term questionnaire references to the interviewing frame, including themes and questions, for this thesis. Nuance of semi-structured interviews is also shown as interviewer’s freedom to ask clarifying and additional questions if needed.

4.2.2. Data analysis

As the data collection is done via combination of structured and semi-structured interviews, the data analysis closely linked to data collection characteristics. When using different themes in interviews, the analysis would be done by thematic reason logically. Themes help the analysis process but the quantitative and qualitative data perspectives have to be considered as the case study strategy contains both perspectives (*see e.g. Ghauri & Grønhaug 2005, p. 113*). For revision, quantitative data is e.g. numeric data or data that can be measured by certain units of measurement and qualitative data is non-numeric and non-quantified data that has to be prepared before analysis. (e.g. Saunders et al. 2009, p. 480).

Qualitative data

Though the qualitative data is diverse and non-standardized, there are some processes that can be used for qualitative data processing (Saunders et al. 2009, p. 490):

1. Summarizing data
2. Categorizing data
3. Structuring data using narrative

Summarizing refers to summarizing or condensation of the important factors, so that the key concepts can be described and presented in fewer words. Categorization refers to development of relevant categories and connecting the relevant data and information to the relevant category. With this approach the results will be divided to smaller entities which on the other hand make the analyzing a bit easier. Structuring with the use of narrative refers to interviewees' stories or narratives which are told subjectively and linked to the real life phenomenon. E.g. when describing experiences or work processes, the use of narrative is effective as the experience or narrative has a beginning, an end and all the action and function in the between. (Saunders et al. 2009, pp. 491-493; 497-498)

Miles & Huberman (1994) present three phases for qualitative analysis (according to Ghauri & Grønhaug 2005, p. 206):

1. Data reduction
2. Data display
3. Conclusion drawing

Data reduction refers to simplification and transformation of the collected data i.e. the raw data is refined to information by adding meaning and sense to it. This phase is quite similar to Saunders et al.'s categorization of the data where categories are built and patterns are discovered. Data display refers to organization and presentation of the data in a way that conclusions and recommendations can be drawn from the displayed data. Again this resembles the summarization of the data, presented by Saunders et al. Finally there is the phase of conclusion drawing, which was discussed in the previous phase. (Ghauri & Grønhaug 2005, pp. 206-207)

As the model of Miles & Huberman (1994) is presented as a holistic entity, the presentation by Saunders et al. (2009) is more like different approaches to adopt for data analysis. In addition to Miles & Huberman's model, Ghauri & Grønhaug (2005, p. 207), present several other analytical activities, that can be used for analysis. Those activities include categorization, abstraction, comparison, dimensionalization, etc. So it is quite obvious as it is stated by Saunders et al. (2009, p. 490): "there is no standardized procedure for analyzing such data".

Quantitative data

Quantitative data is quantified and has units of measurement which makes the analysis much more standardized compared to qualitative data. On its purest, quantitative analysis deals with statistical analysis, variables and terms e.g. standard deviation and correlation. From the perspective of this research, there is no usage or usability for that kind of analysis. Therefore this aspect of quantitative analysis shall be passed and the main quantified data to be collected is the product data content in varying information system.

Suitable methods are e.g. cluster analysis and multidimensional scaling, where cluster analysis focuses on clustering i.e. grouping the units and multidimensional scaling gives an option for observing different variables, e.g. a matrix for product data content comparison between different information systems. (Ghauri & Grønhaug 2005, pp. 198-199) Though the product data content within this research is quantified as in how many data fields there are per information system, the product data contains information about itself, e.g. the data field has a name and a value. Therefore the choice for quantitative method is quite slim as the Saunders et al. (2009, p. 414) describe the raw quantitative data as a data with slight meaning to its interpreter, before it is refined.

4.3. Performing the research

Before the interviews, the interviewees had to be chosen and questionnaires to be created. The research problem is mostly related to ITF, though the parent company, ITG, is relatively important for solving this problem. Therefore, there is a need for interviewees from ITF and ITG. For the questionnaire, relevant themes and topics had to be chosen. The choice of the themes is derived from research problem and objectives. Firstly, there is a need for theme considering the information systems related to product data. Secondly, the importance of product data management perspective cannot be bypassed and the existence of current and required product data has to be solved. Thirdly, the presence of ITG and its norms and regulations has to be noticed. These themes as a foundation for the interview, the interviewees were divided to three distinct categories:

1. End-users from the ITF
2. General issues from the ITF
3. The ITG

Category one consists of employees which are dealing with product data and product data related functions, e.g. information systems, updates and requirements. Category two consists of employees working on a managerial or on an administrative level related to product data and its subareas. Category three consists of employees of the ITG who has the knowledge and experience about ITG regulations, compliance norms, information systems provided by ITG and master data of ITG and its subordinates. The reasoning for this kind of settlement of three distinct interviewing categories is the simplification of the creation process for the questionnaires and the controllability of the interviews.

Selected employees for the interviews are depicted in the table 4. The table 4 includes the identifier (ID), the title of the employee and work description. Also, the interviewees are classified to separate groups, which represent the formerly named category names. References to the interviews are done as follows, e.g. "*interviewee 2C stated that…*" and they refer to the table 4.

Table 4. *Interviewed employees.*

Interviewees			
ID	*Title*	*Work description (IS perspective)*	*Group*
1A	Logistics Coordinator	Product Register, Navision & Kalido (mapping) - main user. Sinfos - maintenance.	ITF - end-users
1B	Sales Support	Siebel & Sinfos - main user. Product Register - basic user	
1C	Finance Assistant	Navision, Product Register, Siebel & Sinfos - basic user	
1D	Summer trainee, Logistics Coordinator	CIS, Navision, Product Register - basic user	
1E	Business Controller	Navision - main user. Product Register - basic user	
1F	Consumer & Brand Marketing Manager OTP	Product Register, KMD - basic user	
1G	Area Sales Manager, South-East	Product Register - basic user	
1H	Head of Sales	Siebel - basic user	
1I	Field Execution Manager	KMD, Kuutio - basic user	
2A	Head of SCM & IS	IS strategy, objectives, operative coordination, budgeting & monitoring	ITF - General issues
2B	System Analyst	Support, design, planning & administration	
2C	Head of Category	Marketing and category management (non IS)	
2D	Market Manager	General management (non IS)	
3A	IS Manager Central Europe North	Define and implement IS strategy for CE North cluster within ITG standards	ITG
3B	Manager Demand Planning & Master Data	Creation and maintenance of local and corporate product master data	
3C	Manager Planning, Data Management & Integration	Technical owner of Kalido MDM and responsible for interfaces between SAP/QAD/APO	
3D	Project Manager Solutions Design SC	IT & Project management for SAP implementations	

As it is shown in the table 4, total of seventeen employees were interviewed for the thesis and the division between categories was as follows: category one, nine employees; category two, four employees; and category three, four employees. For the

ITF, thirteen employees were chosen which presents one third of the whole company's population. Four employees were interviewed from ITG perspective. Interviews as a data collection method, the interviews were held face-to-face obviously, excluding three interviewees from category three, the ITG, as those interviews were dealt via telephone.

After categorization, the construction of questionnaires began. Firstly the main issue was related to the product data content in different information systems. The hindering issue with product data content questions was the usage of the interviews. In the interviewing event, how reliable and valid information would the interviewees capable to give about product data contents without any preparation. Therefore the interviews were held in two sections, where the first section would be filled out in advance and the second section would be interviewed in person.

The interviewing process began with scheduling the interviewing time and date, next the interviewee were provided with questionnaire for section one, then the interviewee answered to questionnaire in written form and sent the questionnaire back to the interviewer. In the official interview, the questionnaire of section one was inspected and clarifying questions were asked for mutual understanding. Then the questions for second section were presented and the answers were documented by the interviewer.

The first section of the data collection is described as a structured theme interview and the second section is described as (semi-)structured theme interview. The first section is called interview instead of survey as it does not provide options for answering, only open ended answers are required. The second section is explicitly an interview and it is semi-structured as it follows certain frame of predetermined questions and it allows the interviewer to use additive questions. The both sections follow certain themes so that is the explanation for the additive term of 'theme' for the data collection methods. The questionnaires for the interviews are presented in appendices 1A – 1F.

While the interviewees were chosen and the questionnaires were constructed, the preliminary study was done. The preliminary study refers to the exploring of the structure and processes of the ITF. This preliminary study provided insight and comprehension of the case company and as deliverables an initial product data architecture representation and theoretical model for product data management was made. The representation of ITF's data architecture is shown in appendix 2. The model for theoretical product data management is shown in appendix 3. The reason for theoretical PDM model is to use it as hypothesis in the interviews and it serves as an ideal solution for PDM on a contrary to the real life situation.

5. RESULTS

In this chapter the results of the data collection are presented. These results are based on the interviews, explained in detail in chapter 4, and to the working period of four months in the case company. First, in chapter 5.1 are the results from the preliminary study, which took place before the interviews. Next, in chapter 5.2 are the results for product data contents per information system. Followed by chapter 5.3, where the renewable information systems are presented and the reasons for the renewals are examined. Finally, in chapter 5.4 are the results of ITF’s current product data management practices and theoretical views of product data management and issues for its development.

5.1. Preliminary study and findings

As a starting point for this research and data collection, a preliminary study was conducted. The preliminary study included exploring the ITF’s products, information systems, processes related to the product data, and the employees and their functioning within product data environment. In chapter 5.1.1 is the brief description of ITF’s products and their structure. In chapter 5.1.2 is ITF’s information systems architecture, which presents the relevant information systems related to product data.

5.1.1. Products

First, this research’s scope was narrowed to the product data perspective and therefore knowledge and familiarity of ITF’s products was crucial. ITF’s selection of goods was examined and it is presented in appendix 4. The selection of goods is divided into four distinct categories: *cigarettes*, *roll-your-own*, *supplies* and *cigars*. Further examination, with the help interviewees 1A & 1B, of the goods revealed the fact that the supplies category contains diverse variety of products. Therefore the supplies category is divided further to five subcategories: *lighters*, *rolling paper*, *filters*, *rolling machines* and *filter tubes*.

Thus, the supplies category is replaced by previous five categories and then total amount of distinct product categories within in ITF is eight. As the research requires examination of product data’s current and future requirements, one sample product was chosen from each category (with the help of interviewee 1A). These sample products are shown in appendix 4, where the products are highlighted.

Second, product structure and hierarchy are quite simplified, which helps the product data management and its development. As the ITF does not manufacture any of its products, the end products are supplied to ITF's warehouse for further distribution to the wholesalers. The end products are distributed to the wholesalers as they were distributed to ITF and the products does not require any revision or assembling. The product structure consists of the casing and its content, e.g. cigarette box and cigarettes. (Interviewees 1A, 1B, 1I & 2A) Though, the product structure depends on the inspected level of product hierarchy. For clarification, see figure 5.1.

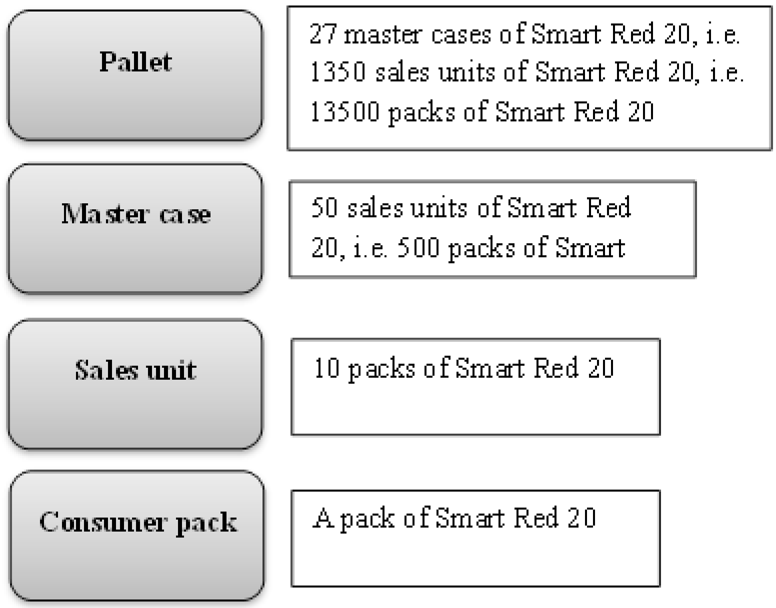


Figure 5.1 *ITF's product hierarchy*

On the bottom level is *consumer pack*, and it is smallest saleable unit. Next is the *sales unit*, which contains certain amount of consumer packs, depending on the product (on the figure, sales unit contains 10 consumer packs). Note that the sales unit is the second and largest saleable unit. (Interviewee 1B) Next is the *master case*, which contains certain amount of sales units, depending on the product (on the figure, master case contains 50 sales units, which is equal to 500 consumer packs). On the top level is the *pallet*, which contains certain amount of master cases, depending on the product (on the figure, pallet contains 27 master cases, which is equal to 1350 sales units, and which is also equal to 13500 consumer packs). (Interviewees 1A, 1B, 1D, 1F & 1G) So, simply said, the product consists of casing and its contents. Also, the products are not composed of any other components or modules. That is, cigarettes cannot be combined with supplies or any other product categories; they are sold as unrelated entities.

Each product, regardless of the level of hierarchy, has certain basic data related to them. ITF uses local SKU codes (LSKU = Local Stock Keeping Unit) to identify its products whereas ITG uses its SSKU and RSKU codes (SSKU = Standard Stock Keeping Unit, RSKU = Retail Stock Keeping Unit) for same purposes (Interviewees 1A, 1B, 1E &

1H). LSKU and RSKU refer to consumer packs that have certain price. E.g. a pack of Smart Red 20 with two different prices would have different LSKUs and RSKUs. Whereas the SSKU would be the same for the both Smart Red 20 packs. Also, products have EAN codes (i.e. barcode number, which will be changed to the GTIN coding), and there are different EAN code for consumer pack, sales unit and master case. Total of three EAN codes per product, which is a lot. (Interviewees 1C, 1G, 1I & 2C)

Furthermore, products have name, content information e.g. how many cigarettes is in the box, details such as size, weight, flavor, country of origin, etc. However, the product data contents per information system are presented in the chapter 5.2 more thoroughly. Beside of the products distributed by the parent company ITG, ITF also sells products from third parties (Interviewees 1A & 2A).

As the ITF going through compliance process, there is a huge variety of identifying numbers (later product number) for ITF's products. There are RSKU and SSKU by ITG, LSKU, EAN codes, and the some customers have their own product numbers. For example, wholesalers TUKO, INEX and Meira Nova have their own specific product numbers as well (Interviewee 1C). Common practice is that the LSKU is used in ITF's organization and, SSKU and RSKU are used when dealt with ITG (Interviewee 1E). EAN codes are used by wholesalers, within sales operations. The specific product numbers for TUKO, INEX and Meira Nova are used within invoicing functions. This may sound disturbing when the orders are made with EAN codes, but the invoicing requires their own coding, which differs from EAN codes. (Interviewee 1A & 1C)

5.1.2. Information systems

As a second deliverable of the preliminary study, information systems architecture from product data perspective was formed (with the help of interviewees 1A, 1B, 1F, 1I, 2A & 2B). Term *data architecture* was used to describe this architectural presentation as it depicts the information systems that are related to the product data and their relations in between. From this point on, term data architecture is used, instead of information systems architecture. Depiction of current data architecture is presented in figure 5.2, and the larger figure is in appendix 2. Data architecture in figure 5.2 depicts the current information systems and their relations to each other. The product data contents for each information system will be presented in the chapter 5.2.

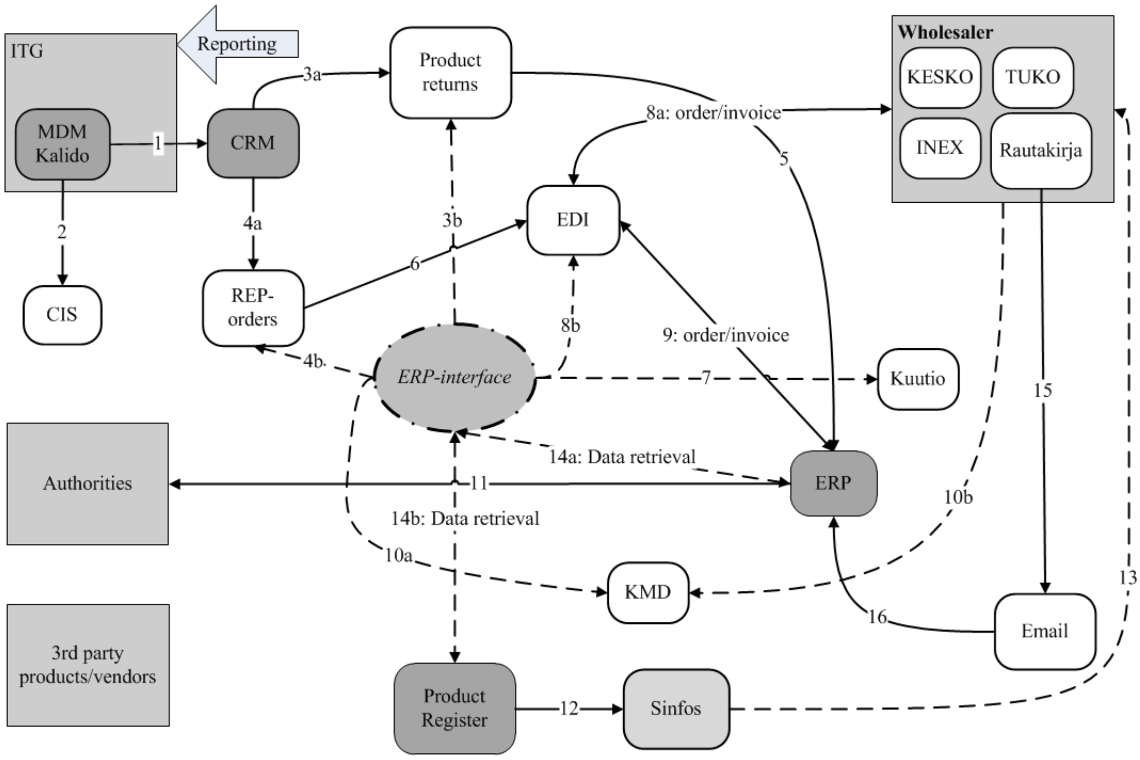


Figure 5.2 *Data architecture for ITF*

In the data architecture, in figure 5.2 and in appendix 2, the information systems are divided into business enabling and business supporting types. Business enabling information systems are those systems that are the main sources of product data and its usage. These systems are depicted in rounded rectangles, with gray colored filling (ERP, CRM, Product Register and MDM Kalido). Business supporting information systems are those systems that use the provided product data in order to support ITF's business processes and functioning (Interviewees 2A & 2B). These systems are depicted in rounded rectangles, with white colored filling (excluding the rectangles within wholesaler entity).

The arrows depict the direction of data and arrows indicated with same number but with different additional signs (a or b) are related to each other. Dashed line depicts information retrieval from specific database or information system. In addition, there is ERP-interface, grey ellipse, which functions as an intermediary between other information systems and ERP's/Product Register's data. Dashed lines 14a and 14b depict data/information retrieval from ERP/Product Register that is done through ERP-interface. Gray colored rectangles depict other parties such as parent company ITG, wholesalers, authorities and, 3rd party vendors and their products.

Data architectures function is to give holistic overview of the current status of ITF's information systems and related parties. It serves as a blue print for information systems renewal and changes as it depicts their relations. Also, it helps with development of product data management, especially with master data.

5.2. Product data in information systems

In this chapter, the product data per information system is presented. The examination of information systems include only the business enabling systems, which are the ERP – Microsoft Dynamics NAV, Product Register, CRM – Siebel and MDM Kalido. Also, the Sinfos information system is examined as it functions as an information system between ITF and the wholesalers, therefore being extremely important part of this examination. The product data contents by each information system are presented in individual appendices, and it also includes presentation of duplicate data fields when compared to other information systems (see appendices 5-9).

5.2.1. ERP system – Microsoft Dynamics NAV

Microsoft Dynamics NAV (*later NAV*) is ITF's enterprise resource planning (*later ERP*) system. It is used for core business functions such as orders and actions related to them, warehouse management and cost accounting (Interviewees 1A, 1E & 2A). The product data input to the NAV is done manually by employees and there are not any data export options[14] (Interviewees 1A &1E). The product data content of NAV is presented in appendix 5. Due to defect of data export options, the product data content in appendix 5 is composed by observing the NAV and with the help of additional screenshots (Interviewee 1A).

5.2.2. Product Register

Product Register contains most of the ITF's product data. In a way it could be called ITF's master data repository for product data. Product Register includes almost all the required data for Sinfos input and it makes the input files for Sinfos (*see chapter 5.2.5*). The product data input to the Product Register is done manually by employees and there are not any data export options, excluding the Sinfos import (Interviewees 1A-1D & 1F-1G). Noteworthy point is that the some of the Product Register's product data is not visible for employees. The product data exists in Product Register's database for Sinfos import, but it can be only examined with administrator rights (Interviewee 2B). The product data content of Product Register is presented in appendix 6, and it only contains the visible data fields. The hidden data fields are presented in appendix 9, where the Sinfos' and Product Register's duplicate data fields are compared. Due to defect of data export options, the product data content in appendix 6 is composed by observing the Product Register and with the help of additional screenshots (Interviewee 1A & 1C).

---

[14] Data export option refers to the exportation of certain data set or entity in readable format. By this is meant a spreadsheet or pdf documentation of the desired data set or entity. Therefore any other export or interface communication is excluded outside of the term *data export*.

5.2.3. CRM system – Siebel

Customer relationship management system, Siebel, is for ITF's customer management obviously and it contains simplified product data content for the sales representatives, so they can make sales orders for stores while visiting in the store. The product data input to the Siebel is done manually by employees and there are not any data export options. (Interviewee 1B) Also, there is a limitation for Siebel that the product data import cannot be done automatically; therefore manual updates are the only option (Interviewees 1B & 3A). For the product data update, the product data source is the combination of Product Register and NAV. The product data content of Siebel is presented in appendix 7. Due to defect of data export options, the product data content in appendix 7 is composed by observing the Siebel and with the help of additional screenshots. (Interviewee 1B)

5.2.4. MDM Kalido

Kalido is ITG's – N.B. parent company's system, not ITF's – information system for master data management. It contains product master data for ITF and currently its relevance is quite minimal, though, the upcoming ERP system renewal will make Kalido's presence more significant and meaningful (for more detail, see chapter 5.3.1). (Interviewees 3B-3D) At the moment, as the ITF has double coding for its products – the LSKU and the RSKU – the products are mapped in Kalido, i.e. ITF's LSKU number is mapped to ITG's RSKU number (Interviewee 1A). The product data input to Kalido is done by the ITG and the data input is initiated due to ITF's new product declaration or change in product specification, e.g. price change (Interviewees 1A, 1F & 2A). The product data content in MDM Kalido is described in the ITG Brand Manual, though a product data exports were used for product data mapping, and the product data content is presented in appendix 8 (Interviewee 3B).

5.2.5. Sinfos

Sinfos information systems is maintained by organization called GS1 in Finland, and globally by the SA2 organization. Sinfos is not ITF's information system, whereas it is developed by the wholesalers in order to provide supplier's information about the saleable products. (GS1) With ITF, Sinfos includes almost all product related data and information that is related to ITF's products. Wholesalers use this product data for e.g. space planning, i.e. for transportation and presentation in the store (Interviewees 1A, 1F, 1I). The product data content in Sinfos depends on the product and the product hierarchy. For example, lighters have certain data attributes that relate to lighters only, e.g. information about flammable gas and dangerous substances.

Also, the hierarchy of product data requirements relates only to the consumer pack, sales unit and master case (*corresponding terms in Sinfos: base item, packaging item and packaging item*). The pallet hierarchy is also included only for its measurement, that is, how much a pallet weighs, and what are its dimensions (width, depth and height). (N.B. this was noticed by examining the Sinfos and its product data content)

Sinfos information can be inputted through interfaces or manually, via internet browser. The ITF used TXT-interface, which in practice means that a zipped folder containing 51 text files is given as an input. (Interviewees 1A, 1B & 2B) Also, these text files have strict names and hierarchical structure for product data import. Unfortunately, the support for TXT-interface was removed at the end of the year 2012 and it will be totally removed by the February 2014. Current and substitutive interface options are XML, Excel and Pricat-messages. The preferred interface for replacing the TXT-interface is the XML-interface, recommended by the GS1 and the most likely option for ITF. (GS1)

At the moment, Sinfos covers most of the ITF's product data content as it stores almost all the data related to the ITF's saleable products (see product data comparison in appendices 5-9). As the product consists of three hierarchical levels – base item, packaking item & packaking item – there are three data object presentations of same product, one by each level. The lowest level – base item – contains the majority of all product data in Sinfos. (Sinfos, by observing) The product data content of Sinfos is shown in the appendix 9 and the product data mapping was done with exported product data files. The product data presentation could be done by hierarchical levels, though the presentation is divided into the categories of Sinfos, so that the product data re-location is easier.

Noteworthy point, by observing the Sinfos, is that the same data field and value can be located is distinct places (i.e. different category or sub category), so the self-duplicates of Sinfos data fields are omitted from the appendix 9. Appearance of appendix 9 differs from appendices 5-8 little, as the appendix 9 depicts the product data sources for Sinfos import. That is, somewhat same as the location of duplicate data, though the appendix 9 includes a column called "SQL". The "SQL" column refers to those product data values, which are not stored in any database rather they are manually coded into the SQL-queries, which create the txt-input files for Sinfos (an example of SQL-query can be seen in appendix 10, where manually coded values are highlighted) (Interviewee 2B). Also, there is the "Notes" column, which provides further information, e.g. is the data value updated manually or is it a combination of Product Register's data value plus some modification in the SQL-query.

5.3. Renewable Information Systems

In this chapter, the renewable information systems and the reasons for the renewal are presented. First, in chapter 5.3.1 is the reasons for ERP renewal from NAV to SAP. Next, in chapter 5.3.2 is the reasons for Product Registers renewal. Finally, in chapter 5.3.3 is the reasons for Sinfos renewal as well as the new product data requirements for Sinfos.

5.3.1. ERP – From NAV to SAP

Currently, ITF is using NAV as its ERP system. Due to the merger with ITG, the ITG has begun compliance processes where it endeavors for consistent processes, including information systems and other IT-resources. Therefore, ITF will be implementing SAP as its new ERP system. The implementation project will be starting in fall 2013. In addition to the compliance reasons, other benefits will become obvious, such as, maintenance, support, license fees, and enhanced reporting. (Interviewees 2A-2D & 3A-3D)

With the implementation of SAP, Kalido's presence becomes more meaningful as there will be interface between Kalido and SAP. Due to the SAP implementation, SAP-Kalido linkage strengthens and it enables data consistency, use of same language for definitions and terms, and it is less error prone for data defects (Interviewee 3B). The master data from Kalido will be imported to SAP, which reduces the manual update actions for product data (Interviewees 3B & 3C).

Overall, the implementable SAP will be as much off-the-self product as it can be, i.e. there will be no customization, except the Finnish legislation requirements. On a downside is the possibility to use MDM/PLM modules in SAP which could be incorporated into ITF's SAP and used for Sinfos purposes, i.e. storing of all product data required by Sinfos and product data imports. Though, this possibility is limited due to budget and ITG's standards. (Interviewees 3A, 3B & 3D)

5.3.2. Product Register

Product Register is an old information system in ITF and basically it has reached the end of its lifespan. At the moment, the Product Register cannot be modified, i.e. no more data fields can be added. (Interviewees 2A & 2B) The alteration of Product Register would be necessary as the product data requirements have increased due to Sinfos version change. Also, the Sinfos input interface will be changed so that the Product Registers core functionality becomes obsolete. (GS1) As the Product Register stores the relevant product data for Sinfos input and produces the input files for Sinfos, it is inevitable that substitute system for Product Register is needed as the current one lacks required data fields and input file producing. (Interviewee 2A)

Apart from Sinfos functionalities of Product Register, the system itself lacks of export option for product data reports and collections (Interviewee 1B, 1C & 1E). In other words, employees requiring and using the product data will have to copy needed product data manually, which increases working time and is not efficient. Therefore the substitute of Product Register would be an information system which would consider changed requirements of Sinfos product data requirements as well as the interface alteration, and the upcoming ERP system renewal should be noted also (Interviewee 2A).

To make matters more complicated, the TXT-file creation for Sinfos input uses SQL-query for retrieving the data from Product Register's database. But, the major concern is the constant values that are coded into the SQL-queries, which means that those data values are not stored in any database (Interviewee 2B). Example of described SQL-query is in appendix 10. Also, the database of Product Register contains some product data values, which are used for Sinfos import, but they are not shown to the users. For example, S-phrases for dangerous goods are in the database but the information is hidden from the regular users, i.e. the employees doing the Sinfos import. (Interviewee 2B)

As the Product Register has to be replaced, the replacement should be done before the SAP implementation is done (Interviewee 2A). The preferred deadline for Product Register replacement is at the end of the year 2013. Therefore, the new Product Register would enable better business functions, i.e. in most cases the Sinfos import, before and after SAP implementation. Also, the new Product Register could store the product data, which could not be stored in SAP.

5.3.3. Sinfos

As it was mentioned before, Sinfos is an information system providing product data for wholesalers (GS1). Though the Sinfos is not ITF's system, it still causes actions for ITF whenever something is changed in Sinfos. Due to changes in Sinfos, the ITF had to take actions. The changes in Sinfos were result from transition of version WS1-Sinfos to WS2-Sinfos. The transition means that the new WS2-Sinfos will be using GDSN-standard (*Global Data Synchronization Network*), which helps suppliers and their product data distribution. With this standard, suppliers are not required to provide their product data in any other product databanks; the WS2-Sinfos will do this instead. This transition created various changes into product data requirements and the timetable of these changes are presented in figure 5.3. (GS1)

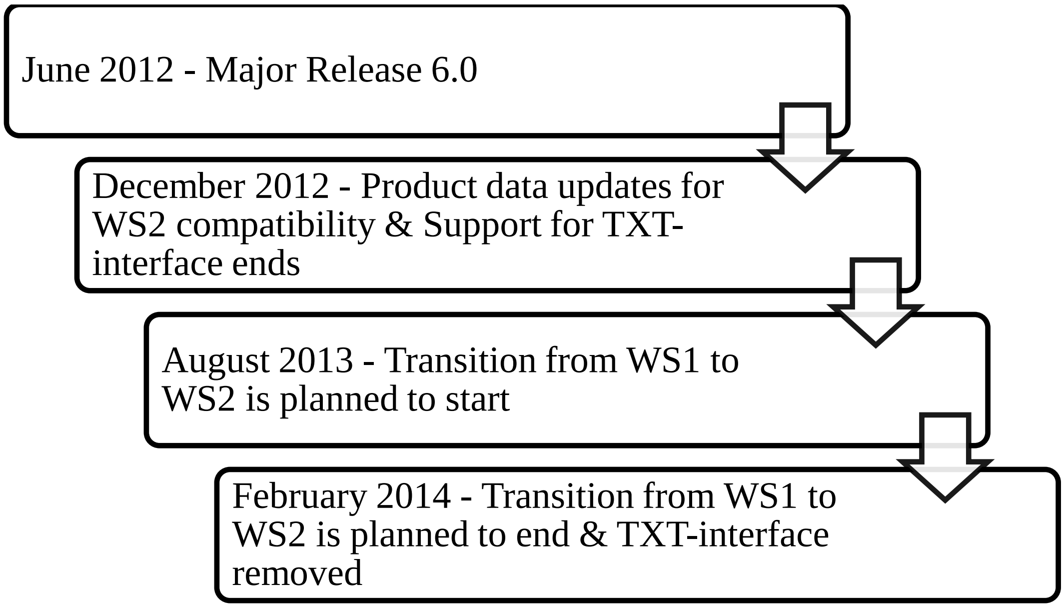


Figure 5.3 *Timetable for Sinfos transition from WS1 to WS2.*

The figure 5.3 emphasizes the key points of WS1/WS2 transition. First, issue was the Major Release 6.0 (*later MR 6.0*), which prepared organizations for the upcoming transition. With MR 6.0, new data fields were added to Sinfos and it forced organizations to update their product data (more about new data fields and requirement in the next chapter). In the next stage, by the end of December 2012, organizations must have had updated the product data. Also, the technical support for TXT-interface was cancelled beyond this point. Next stage will be starting on August 2013, when the transition phase is planned to begin. And finally, in February 2014, the transition should be over and the TXT-interface will be removed, i.e. data cannot be imported through the TXT-interface anymore. (GS1)

Returning to MR 6.0 and its new product data requirements, the MR 6.0 brought three separate data requirements: mandatory data, conditionally mandatory data and optional data. Conditionally mandatory data becomes mandatory when initial requirements are given, e.g. if the product includes within dangerous goods category, provide further information. The new requirements due to MR 6.0 are presented in appendix 11. The appendix 11 contains only those requirements that concern ITF and its products. (GS1)

The publishing of the MR 6.0 made some data fields mandatory, as they were optional in the previous version. Also, some new data fields were added to the MR 6.0, e.g. GPC-codes (*Global Product Classification*), "*trade item unit descriptor*" and "*start availability date time*". GPC-codes relate to product categorization, so that the products of same category can be searched with same code. This code is called Brick-code, and Brick-code for cigarettes is 10000185, so that the search for Brick-code 10000185 lists all cigarettes included in Sinfos. More about GPC-codes and their structure in appendix 12. (GS1)

As the GPC-codes and other new data fields are new to Sinfos, they cannot be found in ITF's information systems'. Moreover, these new data fields can be imported to Sinfos through the TXT-interface, though it has to be done manually, as the information is nowhere to be found (i.e. not in any database). (GS1) Therefore, Sinfos product data updates require manual work as there is no practical reason to create any ad hoc solutions for current Product Register.

5.4. Product data management

In this chapter, the results for product data management and related functions will be presented. First, in chapter 5.4.1 the theoretical perspective of PDM and the current arrangements for ITF's PDM is reviewed. Next, in chapter 5.4.2 the PDM practices in ITF are presented, i.e. what is done and how. Finally, in chapter 5.4.3 the discussion based on theoretical PDM model is presented and its pros and cons are examined.

5.4.1. How PDM is seen?

As the research tries to answer to the question of "How the company should arrange its product data management?", the current state and ideal vision of PDM was examined. When asked about PDM's objectives and its meaning, the answers were quite parallel when compared to each other. Here are some quotes, when asked about PDM objectives and its meaning:

"To offer accurate, diverse and correct information in small amount of applications possible" (Interviewee 1F)

"Continuous process that factors existing and new requirements" (Interviewee 1E)

"PDM is cornerstone of organizations functioning" (Interviewee 1B)

"To ensure timely, correct and cost-efficient storage and transfer of product data" (Interviewee 2C)

"PDM's most important objective is to keep information correct, up to date and available" (Interviewee 2B)

"PDM provides a single data source to deliver product information to all impacted users and systems across our organization" (Interviewee 3A)

As it can be seen, the answers from the interviewees are quite parallel and their vision about PDM is similar, though their areas of expertise and functioning differ (variety of several respondents from distinct categories). The PDM is seen as an important factor within organization's functioning and it is also seen as an organization-wide issue, not just concerning the certain department or business unit. (Interviewees 1B, 1D, 1G, 2D, 3A, 3B & 3C) Despite the uniform perspective of PDM's theoretical ideologies and purposes, there are some issues that rise up during the interviews.

Though, the product data related functions and processes are working, there were some issues considering the PDM arrangement for ITF. For example, interviewee 1I described that PDM in ITF is *"as good as it is possible with these resources"*. This reveals that PDM is functioning, though it could be better. In the list below, there are some issues that were presented problematic with PDM arrangements.

- Too many systems needing manual update actions for synchronization (Interviewees 1A, 1B, 1E, 1I & 2A)
- Product Register does not match current requirements for product data purposes (Interviewees 1A, 1B, 1C, 1E, 1F & 2A)
- Clearer distribution of work and a need for authorized person within PDM (Interviewees 1A, 1F & 1H)
- Existence of duplicate data in various information systems (noted when the product data contents per information system were examined)
- Double coding for products (LSKU/RSKU) (Interviewees 1A, 1C, 1D, 1E & 2A)
- Lack of documentation (Interviewees 1A, 1C, 1D, 1F & 1G)
- Scattered product data (Interviewees 1B, 1D, 1I & 2B)

As it can be seen from the list, there are obvious issues within PDM. On a current status, the product data updates are done manually, which on the other hand increases the possibility of errors (Interviewees 1A, 1B, 1E, 1I & 2A). Product Register is seen obsolete as it cannot handle the current requirements for ITF's product data (Interviewees 1A, 1B, 1C, 1E, 1F & 2A). Duplicate presentations of the same data exist as there are not centralized product database (the Product Register could be considered as one, but it is not capable of function in desired way).

Double codes for product do exist, but the LSKU will be removed due to compliance projects and the ITG's RSKU will be adopted (Interviewees 3A-3D). Final issue deals with distribution of work and need for authorized person, who would be responsible and aware of product data functions and its management. Obviously, this relates crucially to the lack of documentation. As no one is responsible for the product data, then who would be responsible for its documentation. Finally, the product data is scattered across the organization, which reduces the availability of data and the knowledge of data whereabouts is tacit knowledge for each employee (Interviewees 1B, 1D, 1I & 2B). This creates great issues for continuity, for example new employees and their training.

5.4.2. PDM practices in ITF

A need for product data update depends on the changes for product specification, i.e. price, name, etc. Obviously, a new product launch or a removal of a product causes the product data to be updated. Therefore, the product data updates are done when needed, and the amount of update actions is hard to determine or predict. Only in Sinfos, the product data is updated periodically, three times per year. Though, the update dates are well known, it is problematic because ITF has to decide its selection of goods and their prices, four months in advance. (Interviewees 1A & 1B)

In case of any product data related updates, the updates are done manually. And, if the updatable product data has any duplicate presentations of itself, i.e. it is stored in two or more information systems, the update actions have to be done individually to each information system. (Interviewees 1A, 1B, 1E, 1I & 2A) And yes, there are duplicate presentations (or similar) of the same data within different information systems. For example, the description of product may be similar or different, depending on the information system as the updates are done manually. (Interviewees 1A & 1B)

With product data updates, the updated product data sources vary depending on the occasion. The possible data sources are e.g. Consumer & Trade team, Logistics department, Factories, other employees, etc. (Interviewees 1A-1D & 1F) For the business supporting information systems, data sources are the business enabling information systems (i.e. NAV, Siebel, Kalido & Product Register) (Interviewees 2a & 2B). Though, there are concrete data sources for product data, the knowledge of these data sources is tacit knowledge for each employee (discovered by the responses of interviewees). In other words, there are no written documentation about data sources and their data contents.

Also, there are no existing documentation about product data’s whereabouts and contents in distinct information systems, except the Sinfos and MDM Kalido (Interviewees 1A, 1B, 1C, 1D, 1F & 1G). Therefore, the product data mapping was done manually by exploring the information systems, interviewing the interviewees and with the help of screenshots. If the product data content within certain information system is needed, then the only solution is to examine that information system (Interviewees 1A-1C). The documentation for Sinfos is provided by the GS1 and MDM Kalido’s product data content is described in ITG’s Brand Manual, though the usage of Kalido in ITF is quite minimal at this point.

When product data is used for certain functions, the required data is retrieved from the easiest data source available. For example, there are several Excel-spreadsheets that are composed of several product data sources to present needed product data compilation (discovered via interviews). To make matters worse, the number of these spreadsheets is huge, though the complication of the usage of spreadsheets could be solved via proper DMS-system (*Document Management System*), and it is therefore excluded from this research and further examination. At the moment ITF is going through compliance processes in order to be more compliant with its parent company ITG. With these compliance initiatives, the cooperation between ITF and ITG will be better exploited as the Kalido MDM’s potential will be fully harnessed after SAP implementation. (Interviewees 2A-2B & 3A-3D)

5.4.3. Theoretical PDM versus common practice

As a part of preliminary study, a theoretical model for PDM was created. The purpose of theoretical PDM model was to examine interviewees' perspective between theoretical model and common practice. The PDM model was created with the use of theoretical literature and discussion related to the model is presented in chapter 3.1.2, figure 3.2. The PDM model also can be found in appendix 3. The purpose of the theoretical PDM model was to find out that whether the presented model would be appropriate for ITF and what changes should be made in order to make it work. Practically, the founding idea was to receive ideas and thoughts for future PDM arrangement.

Generally, the PDM model was seen functional in a theoretical sense, though the practical implementation would need changes and alteration (Interviewees 1B, 1C, 1E, 2B, 3A & 3C). The model in appendix 3 depicts the PDM system with a single point of entry, so that the product data updates, storing and management is done in one place. Perspective of single point of entry for data updates was seen beneficial as it would exclude the multiple manual updates (Interviewees 1C-1G, 1I, 2A, 2B & 2D).

As the practical benefits were clear for most of the interviewees, the practical implementation of depicted system was unclear and unfeasible in some fashion (Interviewees 1B-1D, 1F, 1H, 2B-2C & 3B-3D). For example, the linkage to the ITG and its systems should be noted (Interviewees 1C, 1E, 2B & 3A) and the Siebel can be updated manually only (Interviewees 1B & 3A). In addition to linkage with ITG; 3rd parties, Sinfos and wholesalers should be considered as well (Interviewees 1E & 2A).

Also, possible future events should be taken into consideration, such as the upcoming SAP implementation (Interviewees 2A, 2B & 3A), as well as the growing use of mobile technology (Interviewee 1I). Furthermore, the expenses for the depicted solution would be too high and it was seen quite ambitious and surreal, in order to work as intended (Interviewees 1B, 1C, 1F, 2B, 3B & 3C). From the data perspective, it was said that the upcoming SAP combined with local MDM system would be sufficient (a local Kalido MDM was suggested, as it is compliant with ITG's standards) (Interviewee 3C). Besides of the functional features and plausibility, the technological perspective was emphasized as well. The upcoming information systems should not be based on certain operating systems, but they should be capable of functioning in different environments and terminals, such as tablets and mobile environments (Interviewees 1D-1F & 1I).

6. DISCUSSION

In this chapter we will discuss the results of this research, presented in chapter five, and present thoughts and give proposals to different issues. First, starting with discussion of ITF’s current state and situation. Then, followed with discussion about ITF’s desired state, including the upcoming future events. Finally, ending with future vision discussion where we will try bridging the gap between current state and desired state.

6.1. Current state

The chapter 6.1 discusses the current state of ITF. The discussion is divided into three different perspectives of product data, information systems and product data management. The chapter 6.1.1 deals with ITF’s product data and its current state. The chapter 6.1.2 deals with ITF’s current information systems and their needs for renewal. The chapter 6.1.3 deals with ITF’s current product data management practices as well as with employees’ perspective towards PDM.

6.1.1. Product data

As it was mentioned as a result of preliminary study, ITF’s selection of goods is quite straight forwarded as the end products do not compose of any other products or equivalent. Consider a regular PC-workstation; it consists of hard drive, CPU, display, etc., so the whole product consists of other products that are also sold separately. In ITF, this kind of situation will not be possible, and this decreases the amount of product related data, i.e. no relations or connections between above mentioned variety pack products, such as PC-workstation. This reduces the complexity of PDM for ITF as the product structure is so simple, which eases the process of PDM development – processes and information systems – e.g. PDM system development with Turban et al.’s (2008; see chapter 2.2.2.) model is reduced in time as the product data related requirements are easy to define, due to simple structure, when compared to the PC-workstation example.

At the moment, according to the interviews, product data is scattered over the organization and there is none or scarce documentation of whereabouts of product data and its content. For example, some product data may be located in a network server drive, which may be publicly available or in someone’s private drive. Furthermore, are the employees aware of this information, and if they are, will they remember location of that data in question? From viewpoint of PDM/MDM (chapters 3.1 and 3.2), it is crucial that the organization knows its own data and its location, and information possession.

For the ITF, it is a matter of opinion that should we speak about PDM initiatives or MDM initiatives. That is, PDM covers all the product related data and functions within organization (CIMdata 1997; Kropsu-Vehkapera et al. 2009), whereas the MDM encompasses only the business critical data objects, therefore being more limited than PDM (see Loshin 2009). Though, as the Sinfos has the largest product data content (see appendix 9, compared to the appendices 5-8) and is vital information system for ITF's existence, as it functions as an interface between wholesalers, therefore enabling sales and business making. So, it could be asked that is the product data content in Sinfos considered as master data or where the line is drawn?

There is no correct answer for ITF's product master data, as the response will vary on the respondent. Though, some outlining can be done and it requires the examination of business enabling information systems: NAV, Product Register, Siebel and Kalido, and the Sinfos. First, MDM Kalido's current linkage to ITF's functioning is so minimal, that it can be excluded. Second, product data in Siebel is originally from NAV and Product Register, therefore Siebel should be excluded as well. Obviously, the current ERP system, NAV, is important for ITF's master data. Distinction between Product Register and Sinfos may sound vague, but the Sinfos' product data content is more relevant than Product Register's as the Product Register does not contain all the data for Sinfos import. Therefore, the ITF's product master data could be described as product content within NAV and Sinfos (see appendices 5 & 9).

Before MR 6.0 release for Sinfos in June 2012, ITF was capable of dealing with its product data. After the MR 6.0, ITF is still capable of functioning, though it requires more work as the product data requirements for Sinfos have increased. MR 6.0 brought new data attributes and fields, which are mandatory in Sinfos in some extend. And the novelty of these data requirements changes the status quo, as the ITF cannot store these data attributes in its information systems.

The obvious information system would be the Product Register, but it is old fashioned and uses aged technology, and therefore it cannot be altered to cover the new requirements. Surely, it is possible to alter and modify Product Register if required, though it is made in the mid 1990's and it has been under alteration several times, so the most lucrative option, financially and technologically speaking, is to replace Product Register with other solution. In addition the theoretical literature for information systems renewal (chapter 2.2.2) speaks up for this arrangement, as the Product Register has exceeded its lifespan and is incapable of fulfilling its purpose of use.

Although the new requirements for Sinfos are given in precise detail, some of its determination is not fully complete. This concerns the new GPC system, where product can be classified into distinct categories, and this helps the wholesalers as they can view all certain products within certain category. E.g. all cigars, which includes all the cigar products from every manufacturer.

The structure of GPC system is depicted in appendix 12. On the highest level is the *segment*, then comes *family* and after this comes *class*. Under class-level are *brick codes* that have *attributes* and *attribute values*. Noteworthy point is, that brick codes are the highest level included in Sinfos, i.e. only brick codes, attributes and attribute values are in Sinfos. Though, the higher levels of GPC hierarchy are excluded from Sinfos, they should be included in the new Product Register in case those hierarchy levels will become part of Sinfos. Also, the brick code attributes and attribute values are optional in Sinfos, though they should be taken into consideration as they will become mandatory after the WS2 transition.

Returning to GPC system and its determination, there is a minor defect on it. For example, brick: *cigars,* includes attribute: *cigar style* and the possible attribute values are: *extra long, regular, unclassified* and *unidentified*. The problem is that nowhere is determination for difference between extra long and regular cigar. Since summer 2012, there has been no change to this matter (and now it is beginning of the year 2013) although there has been discussion about this flaw with GS1 back in the 2012. This dilemma creates challenges for ITF's PDM initiatives as the GPC-system is the basis for the change, though GPC itself is incomplete as a classification system.

6.1.2. Information systems

Current status with information systems is fine as everything is functioning and nothing is broken. Though, this does not mean that nothing should be done. On the contrary, technologies change and information systems evolve all the time, as it is stated in the theoretical section. With ITF, compliance processes require that the ERP system should be changed to SAP. Also, the Product Register is becoming obsolete due to Sinfos requirement changes. Apart from renewable information systems, Siebel does not require any changes or alteration. Though, it lacks automatic data update option, which lessens possible options for ideal information systems architecture and product data management. As the theoretical sections describes, the more the information systems can update data automatically, the better. Also, alike Siebel, MDM Kalido does not require any changes, i.e. changes cannot be made, because it is ITG's system. Unless there would be extremely business critical reasons for ITF's functioning, then possible alterations could be done. Due to upcoming SAP implementation, the presence of Kalido becomes more relevant as it provides master data updates to SAP. The relevance of SAP will be discussed further on the upcoming chapters 6.2 and 6.3.

As the new requirements and changes for Sinfos WS2 were introduced, the ITF had to take action and they have seen that the Product Register have become obsolete, as its core functionality relates to product data import to Sinfos. The Product Register cannot fulfill its purpose anymore and it is lacking other important features as well, such as data export for further examination and usage.

ITF is currently using TXT-interface for data import and there has been no maintenance or support for TXT-interface since the beginning of the year 2013. After February 2014, the TXT-interface is no longer available and ITF has to adapt new interface option for product data import. This option will be XML-interface as suggested by GS1 and generally chosen option of ITF, according to the interviews. Apart from the usage of interfaces for Sinfos import, the import could be done manually via web browser. Though, the manual updates would end up in same results as the import via interface, but that choice is inefficient and will be excluded from the set of possible solutions.

The GS1 has suggested that the transition process, especially resigning of the TXT-interface and adoption of the new interface, should be started in good time, i.e. before the February 2014. As the option for new interface have been agreed, then the emphasis of Sinfos and Product Register linkage is moved onto the data sources for the new interface. Currently, Product Register provides almost all data for Sinfos, but the new requirements have made Product Registers actions insufficient. Therefore, encouraged by the interviews and theoretical section, ITF needs a replacing solution for Product Register, which could encompass the new requirements as well as the old contents of Product Register, and include desired features and functions, and be capable for future development. For now, ITF has to use TXT-interface and add the missing data fields manually.

Some of the product data is incorporated into SQL-queries, which create the txt-files for Sinfos import. This incorporated product data is static data within the SQL-queries and it is a horrible situation from the developers point. There are thirteen distinct txt-files, where static data is used in SQL-queries. Each of them files has same data attribute, an identification number for manufacturer or in this case, data provider, i.e. ITF. Consider the situation where the identification number in question would change. The change would have to be done in each file separately. Though, this does not sound were difficult, but there are easier ways to do this. And this kind of operation requires the administrator for incorporating the changes. If it would be a data field in an information system, a regular employee with access rights could change that data attributes value.

Although this sounds as shocking example of software engineering gone wrong, there is a little sense in this madness. For what comes to Sinfos, it requires certain data values, which will present values yes or no, as: is this product under the alcohol taxation? And there are many similar data fields in Sinfos, so it is up to the developer that is this information static in SQL-queries or is stored as a data field in information system. Either the case, theoretical section suggests proper documentation that will depict the chosen arrangement. And then, the information should be made available for distribution as it is the key point of knowledge management (see for chapter 2.1). In this case and currently within ITF, if the meaning of every static data value in SQL-queries has to be known, a TXT-interface manual shall be revised.

Lastly, the information systems and their relations seemed to be functioning, but an actual case from ITF proved that a single missing data value can lead to missed orders. The missing data value was an RSKU number *X* from NAV. The missed orders occurred when sales representatives made orders. The normal routine of rep order is that the order is done with Siebel CRM and then processed in REP-orders application.

Then, the REP-orders application retrieves supplementary product data via ERP-interface and then sends the order into EDI-application (*External Data Interchange*). From there the order is transferred to NAV and collected in the warehouse and so on. The problem occurs, when supplementary data is retrieved via ERP-interface for *product_X*. The query tries to find supplementary data from NAV with RSKU value *X*. As the RSKU value *X* is nowhere in NAV to be found, the query results with no supplementary data. This leads to zero-line in the EDI and the order for *product_X* does not transfer to NAV.

Overall, this error was noted due to lucky coincidence. The sales rep was struggling with his or her order, so the order was done by another employee, who is responsible for rep-orders. This employee found that the certain product in that order did not transfer to the EDI as intended and the problem was solved after this. Financially speaking, this was not a major setback though it gave a remarkable lesson for data quality assurance. Then, why that RSKU number *X* was missing from NAV? Answer is quite simple as humane error and the fact that the NAV does not have a programmatic check for missing data values. That is, the RSKU value is not a mandatory value programmatically, but in a business sense, it is.

6.1.3. Product data management

Theoretically speaking, the purpose and function of PDM were seen clearly, almost like a model answer for an examination as the answer were alike the theoretical literature of PDM (see chapter 3.1). Though, most of the interviewees emphasized the fact that things could be done otherwise and there is room for improvement. For example, there were no clear picture of PDM practices and arrangements in ITF. Therefore, clear distribution for PDM related work was needed, authorized person for PDM suggested and enhanced documentation was wished for, alike as the theoretical PDM literature suggests.

Currently there is a segregation of duties for certain product data related events and actions. For example, one is responsible for PID process in case of new product or product changes and the other is responsible for product data updates. Although, the certain roles and responsibilities exist, as an outsider (researcher), there was no information available that who did what and when. Basically, by asking this could be solved, but this is not a proper way according to the theoretical literature of PDM or knowledge management.

On the current basis, employees know what to do, how to do it and where to find information if needed. This quite briefly summarizes the present situation within ITF. For example, a case of product data update: *an employee A retrieves updated product information from data source X and does the required changes and updates into information system J. After this, employee A informs the employee B that the product data needs update actions. The employee B updates the information system K and L.*

This is a working solution as long as nothing changes drastically, e.g. employees. But, from the concept of PDM and knowledge management, this is unbearable situation and needs at least some documentation of these processes and its steps. In addition with knowledge management concepts, the sharing and distribution of knowledge should be emphasized as well (see chapter 2.1).

As the product data processes and actions are what they are, the product data related updates and changes are done when needed, depending on the situation. The usual case is when something in the product's specification changes, usually the price. Also, the Sinfos requires periodical product data updates, consisting of three updates per year. The problem of Sinfos updates is that the updates are done in four month periods and four months in advance, so that the product data update in January, will concern products and their information from May to September. So, the ITF will have to confirm their product prices and other details in advance, before the products are supplied to wholesalers. This creates a situation where assumptions have to be made and this requires knowledge management and supportive PDM for making the right decisions (see chapters 2.1 and 3.1).

The comparison between theoretical PDM model and current situation revealed some issues that should be solved. The main issue is the manual update actions and the usage of product data. First, as the case example – given in the chapter 6.1.2 – revealed the vulnerability of ITF's PDM practices. Only one missing data value was needed for this case. Though the error was discovered by fortunate coincidence, it should be clear that this kind of occasion should not be possible in the future. Second, the usage of product data was seen defective as the information systems lacked of proper export options and data processing tools. From information systems literature perspective, this leads to situation where the information systems do not fully enable business as intended (see chapter 2.2). A good example of badly functioning information system is the Product Register, where the required product data has to be manually copied for other purposes, excluding the Sinfos import. Also, related to Product Register, it was mentioned that the shortage of RSKU code from the system was seen as a flaw. And further, the lack of RSKU code in Product Register has led to point, where the required RSKU code is retrieved from other data source, which is e.g. Excel spreadsheet. Therefore, the employee's work is based on a secondary or even tertiary data source, which is total opposite of knowledge management, PDM or MDM concepts (see chapters 2.1, 3.1 & 3.2).

6.2. Desired state

The desired state of ITF is discussed in this chapter. The discussion is divided into three perspectives; product data, information systems and product data management. The desired needs, requirements and wishes are compiled from the interviews.

Product data

Generally, the employees of ITF emphasized the importance of product data quality, alike the PDM and MDM concepts do (see chapters 3.1 and 3.2). With the term quality, issues as reliability, correctness, availability, position for product data were mentioned during the interviews. The product data should be reliable in a way that it is up to date and it is correct, regardless of the data source (opinion mutually agreed between interviewees and theoretical section). Product data should be correct so that it can be used without checking; also the programmatic check would be great, as the case example depicted (see chapter 6.1.2).

Availability relates to information systems and user rights, as the use of information systems is restricted. Therefore, the availability of required product data for employees has to be ensured. Otherwise the needed product data may be retrieved from poor quality data source, which is against PDM, MDM and knowledge management theories (see chapters 2.1, 3.1, 3.2). Position of product data relates to the availability perspective, as the employees need to know that where to get certain information. This issue is solvable by documentation and knowledge sharing as depicted in theoretical section (see chapter 2.1)

Desirably, the centralization of product data was seen as an option for product data quality improvement. If all the product data would be located in one location, a single database or similar solution, then the whereabouts of product data would be solved and the availability would be easier to arrange by granting read-only access rights to the system. This centralization concept is similar to the MDM's architecture types (see chapter 3.2.3). The centralization solution is seen more ideal as the product data is currently scattered around the ITF, though, with sufficient documentation the position issues could be solved (see chapter 2.1), i.e. whereabouts of product data per information system. Apart from position and availability issues, the reliability and correctness of product data relate to the PDM activities and processes, not solely to the technological solutions, i.e. information systems (see chapter 3.1).

Information systems

From the technological and information systems perspective (practical and theoretical viewpoints included), it was desired that the current Product Register would be replaced with another solution. The most important reason for replacing Product Register is the new requirements for Sinfos, i.e. the removal of TXT-interface and new product data requirements. Also, the Product Register has exceeded its lifespan and therefore cannot be altered for future challenges. So, from the lifespan perspective, Haikala & Märijärvi (2006) state that the information system should be removed from usage or replaced as it has reached its lifespan, i.e. it is obsolete or is based on old technology, alike the Product Register. The Product Register replacement is desired to be implemented before expiration of TXT-interface and preferably before SAP implementation, in fall 2013.

The next inevitable step is to change current ERP system NAV to ITG's SAP system. The implementation of SAP increases the level of compliance between ITF and SAP, as it is remarked from the interviews (see chapter 5.3.1). Also, due to this SAP implementation, the MDM Kalido can be exploited for master data import from Kalido to SAP. This decreases the amount of manual data update, which is a greatly beneficial issue. (See chapters 5.2.4 and 5.3.1) As the SAP implementation is unavoidable, the combination of SAP and the substitute of Product Register are seen as a desired foundation for ITF's PDM arrangements and functions.

Overall, the most desired feature for holistic information system management and functioning was the synchronization of the information systems. By this, the information systems would share and distribute their data back and forth, which would lead to decrease of product data related updates. Yet, the need for improved data analysis and processing tools were required as well, mostly related to the Product Register, which is on the renewable information systems list. Some technological desires were presented, as the use and exploitation of mobile technology and devices. Also, the platform independence was suggested as well. (See chapter 5.4.3)

Product data management

For the PDM and related functions, the segregation of duties should be done clearer and documented for future events. Generally, the importance of documentation, that is lack of it, was emphasized and it is therefore crucially desired. Both the interviews and theoretical section mutually agree on the importance of documentation. With clear segregation duties, everyone will know their function and the documentation supports this knowledge.

Also, the problem of redundancy, in product data as in business processes, should be avoided as well. According to the concept of PDM, the amount of manual work should be lesser due to the synchronization and automated processes (see Kropsu-Vehkapera et al. 2009). The lessening of manual update actions related to product data was seen desired, as at the same time it would decrease the possibility of humane error (see chapter 6.1.2, the case example of humane error).

Holistically, the most desired function for PDM and overall business functioning was the importance of knowledge management. At the moment the knowledge is dispersed across the organization and most of the knowledge is tacit knowledge within employees (see Nonaka & Takeuchi 1995). Therefore the externalization of tacit knowledge and its dissemination is seen crucial for ITF and its development (see chapter 2.1.2; Nonaka & Takeuchi 1995). Also, the transparency within PDM functions should be accomplished as the PDM function should be public, not hidden and vague as it is now. Though, the documentation and new guidelines are not effective, unless employees are trained and instructed.

At the current moment, the ITF has several identification numbers for products. This is quite disturbing and may lead to errors, but due to the compliance projects and SAP implementation, the LSKU numbers should be removed, and RSKU and SSKU numbers are adopted for primary identification (and this is a wish from the employees). Yet, the variety of different codes for each wholesaler is more complicated, as the issue cannot be decided by the ITF. Still, the removal one excess product code is better than implementation of another.

6.3. Future vision

In this chapter the distinction between the current state and the desired state is examined. The discussion of future vision is divided into two chapters, where chapter 6.3.1 discusses the future state of product data and product data management, and the chapter 6.3.2 discusses the future state of information systems and their renewal projects in ITF. Basically the founding idea is to find the missing links between current and desired state and present the suggested actions for ITF's future activity.

6.3.1. Product data management

The desired features for product data and its quality are quite concrete and plausible into certain extend. The development of quality features for product data can be difficult and requires a lot of work, but the achieved results will pay off in the long run and help the development and functioning of knowledge management, PDM and MDM initiatives (see chapters 2.1, 3.1 and 3.2). Obviously, without development and improvement, the employees will work as usual, wasting precious time to the detailed issues and have to endure uncertainty of reliability in their daily routines. The data quality improvement requires detailed specification of what should be done and what is achieved with it. Alike with knowledge management and information systems renewal (chapters 2.1 and 2.2), they need measurement and management. Measurement is required to keep track of the changes between old and the desired state. Management requires that the schedule and the specification are followed and it is also responsible for measurement actions.

Key concept of knowledge management is that the knowledge should be shared and disseminated across the organization, which will lead to individual learning and at the end, the whole organization learns (see Nonaka & Takeuchi 1995). As the product data and its functions are not relevant for every employee in same extend, the importance of available information for these issues should be everyone's concern. Though the current employees know what has to be known and they will know where to get the desired information, but what should be done in case of a new employee. In SECI-model (figure 2.3, p. 16), the socialization describes the situation where the new employee learns by observing the old employees. This is useful for a while, but the new employee needs the other three knowledge conversion processes for developing his or hers intellectual assets, that is tacit knowledge. For example, by examining process charts and other documentation, the new employee internalizes explicit knowledge that in time will turn to his or hers tacit knowledge. (Nonaka & Takeuchi 1995)

Currently, the ITF's product data is scattered across the organization and in ideal situation, the product data would be centralized into a single data repository (practically desired but theoretically unusual, see chapter 2.2.4). First, the product data has to be mapped out, as it is done in the appendices 5-9, and extracted from the current information systems.

Noteworthy point is to notify MDM Kalido's product data content, as its presence comes more evident for ITF after the SAP implementation. Then, the principles of data migration should be applied, i.e. duplicate presentations should be removed and data should be cleansed. Finally, the transformed data should be loaded into the new data source system. (Thalheim & Wang 2013, see chapter 2.2.4) Although, the results of the interviews showed that the single data repository idea is not plausible solution, the data migration process should be executed, but with significant difference of the load process. On this case, the transformed product data will be loaded into several information systems as the single system model is excluded.

The most suitable option is to harness the resources of upcoming SAP and *new Product Register* in order to store the product data, specially the product master data. It is hard to define ITF's product master data as master data is considered as business critical data. For example, the product data in Sinfos is business critical, but should it all be considered as master data? (See Loshin 2009) Therefore, the distinction between product master data and product data is quite complex. Anyway, if the *new Product Register* would contain only product master data, then were the other supplementary product data would be stored? Thus, it should be clear that the "all" product – consisting of product master data and supplementary product data – data should be stored in a same information system, which is the *new Product Register*.

After the product data has gone through the data migration, the practices and arrangements for PDM should be implemented. The development of product data and its management are closely linked together. Thus, proceeding parallel with data migration in order to ensure the best possible practices. (See Thalheim & Wang 2013) The best PDM practices should include the determination of people in charge of PDM and related functions. Also, the practices should include clear segregation of duties and responsibilities. Lastly, all the PDM practices and functions should be documented for further use and revision. The documentation should encompass the structure of PDM; the data sources, information systems, employees, etc. (See chapter 2.1 for knowledge management)

Due to the compliance process, the ITF will be abandoning its LSKU codes and adapting the ITG's SSKU and RSKU codes. The alteration from code to another seems quite simple and straight forwarded, but it contains huge possibility for irregularity in ITF's business functioning. That is, as the ITF has used its own LSKU for reporting and other purposes, it has to track those processes and actions and reconsider the possible effects of RSKU adoption. So, how the product code changes affects ITF's business and functions. Abandoning of LSKU should be done via data migration; though the possible challenges should be foreseen, e.g. migration challenges chapter 2.2.4, p. 33 (Kelly & Nelms 2003). The double coding could be maintained via different technological solutions, but it would be easier to adapt the new product codes immediately than postpone the full adaption.

6.3.2. Information systems

As it has been stated, the ITF is inevitably facing information system renewals and changes. The most important factor for these information system renewals is the schedule for each event. The outline for information systems changes and renewals are depicted in figure 6.1.

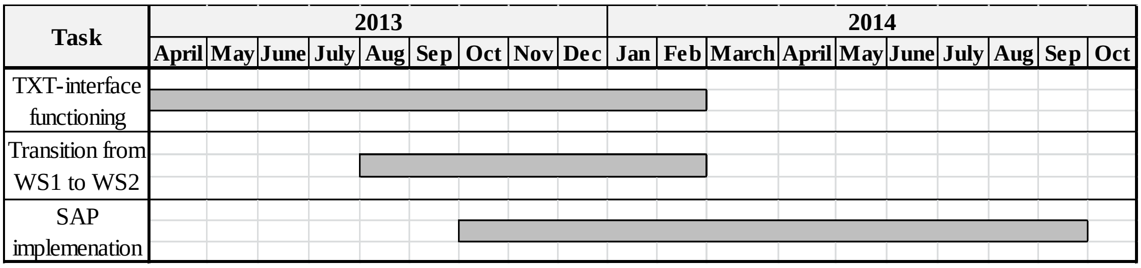


Fig. 6.1 *Schedule for information systems renewal and known deadlines*

As it can be seen from figure 6.1, the changes and renewals are parallel to each other. This creates some limitations for possible solutions as the processes are executed at the same time. First off all, the Sinfos transition from WS1 to WS2 is scheduled from August 2013 to February 2014. At the same time, the TXT-interface for Sinfos is still in use, though it does not have any support and it will be removed after February 2014. Meanwhile, in October 2013, the SAP implementation is scheduled to begin, and it is planned to end in October next year. The wish from GS1 was that the abandoning the TXT-interface should be done beforehand and preferably before the February 2014. If ITF adopts the new interface in good time before the deadline, most likely the XML-interface, then the combination of SAP and the *new Product Register* cannot be used to the product data import purposes, as the SAP implementation will be under execution.

Therefore, the *new Product Register* has to be able to contain the required product data for Sinfos. Though, it should be noted that the known basic product data within SAP, will be imported from Kalido (see chapter 5.3.1), and some of the product data in Kalido would be useful for Sinfos import. But the presentation of that data is different and therefore useless, i.e. would require static conversions and increases maintenance costs. For the SAP, it will contain RSKU and SSKU codes, though those codes are unnecessary for Sinfos functioning, as the EAN codes (replaced with GTIN codes) are used. Thus, the linkage between *new Product Register* and SAP should be taken into consideration (e.g. see Turban et al.'s (2008) stages for information system renewal process, in chapter 2.2.2). For example, in case of reporting the report may need supplementary data from *new Product Register* and there has to be some relation or connection between SAP's and *new Product Register's* product data objects.

So, from the time perspective, the acquisition of *new Product Register* has to be done as soon as possible, preferably before August 2013, when the WS1 to WS2 transition begins. One reason in favor of this settlement is the new XML-interface for Sinfos. To adapt the new XML-interface, ITF needs the new replacement for the old Product Register. Then, the ITF can develop and test the new XML-interface with the *new Product Register*, and the old Product Register can be held as backup solution, providing the files for TXT-interface. The process could adapt the waterfall model by Haikala & Märijärvi (2006; see figure 2.5, p. 22), which is designed for IS development. With this preparation, there should be enough time for adapting the XML-interface in order to abandon the TXT-interface. As mentioned before, the *new Product Register* should be able to interact with SAP, so this has to be noticed, when the specifications are made (see Berghout et al. 2011; Turban et al. 2008).

The *new Product Register* should encompass the following features and functionalities: product data export options, i.e. various import options such as Excel, txt-file, xml-file; tools for product data processing, creation of xml-files for Sinfos import; interface to SAP; programmatic data check functions, in case of data addition or modification; and simple user interface for usability reasons. Furthermore, the *new Product Register* should be flexible enough that is can be modified in the future. The given listing is not exclusionary, and it presents the required and desired features and functionalities, which were perceived during the interviews.

To make the *new Product Register* acquisition more complicated, the acquired information system or solution shall be compliant with ITG's technical standards and guidelines. For example, ITF could adopt its own local MDM Kalido and use it as substitute for the old Product Register (see chapter 5.4.3). Moreover, it is known that the Kalido and SAP have interface between them and this could be drastically facilitating solution (see chapter 5.3.1). Though, as it would be compliant solution for ITG's standards, the practicality of it has to be discovered as well as the functional perspective.

Returning to the SAP implementation, the renewal of old Product Register has its impact on it. The key thing is to define and create *new Product Register* as flexible as possible, so it could operate with SAP in the future. Whereas the SAP itself contains MDM or PLM modules, which could be used for Sinfos purposes. Though, it is outside the conversation as it would be expensive and not compliant with ITG's standards. In addition, what would be done in the mean time when TXT-interface is removed and ITF does not have SAP yet? So, the product master data inclusion within SAP is excluded from possible solutions list. Apart from Sinfos and Product Register renewal, the SAP requires accurate definition for its implementation, although the SAP will be customized as little as possible. In this way, the implementation will be more meaningful and the avoidance of possible alteration work after the implementation is direct savings expenditures (see Haikala & Märijärvi 2006).

7. CONCLUSIONS

In this chapter are the conclusions for performed research and study. The chapter 7.1 summarizes the found results and given recommendations for the case company. The chapter 7.2 contains assessment of the study, contributions and discussion for further research.

7.1. Summary of results and recommendations

This chapter presents a summary of the results and the given recommendations. Chapter 7.1.1 summarizes the findings of this study by starting with the theme of ITF's product and continuing with information systems, product data content, renewable information systems and ending with product data management. Chapter 7.1.2 presents the given recommendation for ITF and the recommendations are given thematically regarding to product data, product data management, and information systems and their renewal. And finally, the chapter answers for the research questions, presented in the chapter 1.2.

7.1.1. Summary of results

First, ITF's selection of goods is not too broad, consisting of four major categories. Within these categories the supplies category includes variation between products, and from the viewpoint of product data, the supplies category can be divided into five separate categories. Therefore, the final amount of product categories is eight. For the structure of products, the composition of products is simple alike the selection of goods. The products are composed of packaging and the content of that package. There are no variety packs or similar, where the product would compose of other products, e.g. no product variety pack for lighter and cigarettes combination. The lighter and the cigarette pack are sold as separate products.

Second, ITF's information systems and their relations to each other. The information systems can be divided into business enabling and business supporting systems. Business enabling systems – NAV, Siebel, Product Register and Kalido – are systems which are crucial for ITF's functioning. The rest, business supporting system, are system which support and diversify ITF's business making. For clarification, Sinfos is crucial component of ITF's business, though it is excluded out of the business enabling solutions as it is not ITF's information system. Therefore examination of ITF's product data is focused on the business enabling information systems, and Sinfos obviously as it is inevitable for ITF's business making.

Third, product data content in each information system varies and the most important information systems for ITF are the NAV and Sinfos. Siebel's product data is from the NAV and Product Register. Kalido's significance will increase after the SAP implementation and in the meantime it serves as mapping tool between ITF's and ITG's product codes (LSKU – RSKU). Product Register is the data source for Sinfos import and it covers almost every data field required by Sinfos. Because of this, the product data content in Sinfos covers Product Register and even more. In combination with NAV and Sinfos, the product data content can be described as ITF's product master data.

Fourth, renewable information system for ITF are NAV, Sinfos and Product Register. The NAV will be replaced by SAP and this is done due to the compliance and compatibility issues, initiated by ITG. Sinfos' data requirements and interface, have and will change, so the ITF have to answer for the changing needs of Sinfos. The TXT-interface will be removed and XML-interface will be adopted. For the new data requirements, the Product Register would be the answer, though it is out-of-date and cannot be modified according to the new requirements. Therefore, the Product Register should be replaced with a new solution, which can handle the new requirements, communicate with SAP and it should be flexible, considering the possible changes in the future. The renewal process of information systems will not be easy as the renewal schedules are overlapping. The SAP implementation and the Sinfos transition from WS1 to WS2 are parallel to each other, and it would delightful to have the new Product Register before this.

Fifth, the product data management within ITF. The product data is scattered across the organization and there is scarce or absolutely no documentation for product data or product data related processes. The knowledge of product data, its whereabouts and related processes are mostly concealed within employees as tacit knowledge. The employees in ITF are conscious about this and they have stated that it could be done otherwise. Though, it was expressed that everything is fine within given resources. Overall, the current PDM requires lot of manual work, which increases the possibility of errors (e.g. the case of missing RSKU; *see chapter 6.1.2*). The amount of manual work is the result of poorly connected information systems, as the each information system has to be updated manually. Also, the update actions include updating the same data in several information systems. This leads to the situation of duplicate data, which is not desired from the PDM or MDM perspective.

All things considered, the situation in ITF may sound desperate and even hopeless, but on the bright side, everything is functioning and ITF is performing its business. Though, it should be noted that something could be done and moreover, should be done. By the given recommendations in the next chapter, ITF can reduce some of its excessive and overlapping work functions and improve its effectiveness.

7.1.2. Recommendations

"If it is not broken, do not fix it" is a saying that can be heard every so often. This may hold up, but when the great "what if" scenario happens, then what will be done? With "what if" scenario is referred to the actions and occasions that may lead to disturbance in the organization and its functioning, i.e. business making. For example, what if the person(s) in charge of business critical functions will be prevented, e.g. accident or similar, then who will be in charge and will the stand-in person be able to deliver desired and required results?

As it was mentioned earlier, in the chapter 7.1.1 that currently ITF is capable of functioning as intended from the business perspective. Though, the ITF had some issues, which need to be resolved. As the issues are not problems of tomorrow, they will become ones, if nothing is done in advance. Recommended improvements and adjustments relate to ITF's product data and its management, information systems and their renewal, and overall administrative factors within product data perspective.

First, the product data in ITF should go through revision process, where the consistent definitions and terms for product data are created. For example, the naming convention of products, how it is done and what is the structure of the product. As an example, Siebel's product description includes the price of the product, whereas the price is excluded from the description/name field of Product Register. The consistent structure for product names and definitions helps the other employees when they know which search words should be used for product search. As the ITF possess a distinct variety of product identification numbers, it would meaningful to remove the excessive codes. The upcoming SAP implementation creates the opportunity to improve processes in order to remove the LSKUs and adopt the RSKU's by ITG.

Second, the product data management including data whereabouts and update processes. For the moment, most of the relevant and necessary information and knowledge is embodied as employees' tacit knowledge. Therefore, there is scarce explicit knowledge, e.g. documentation, of product data processes and product data whereabouts. This creates discontinuity for ITF in case the employees change. In case of employee replacement, new employee(s) have to learn everything by doing and internalize absorbed knowledge as their tacit knowledge. This leads to unnecessary work, from which nobody benefits. So, the significance of documentation is highlighted to ease the work of future employees. Moreover, it was desired that there would be a person in charge of PDM, who could help other employees with PDM related questions and functions. Also, the segregation of duties should be made more transparent, so that others would know that who is responsible for what. All this requires education and employee training, which is supported with documentation that depicts the chosen arrangements.

Third, the product data needs to be mapped and migrated for the upcoming information systems renewals. First renewal is to replacement of the current Product Register with a new one that is capable of storing all the relevant data for Sinfos, is flexible for future needs and can communicate with SAP via interface. The product data mapping is done and the product data contents are presented in the appendices 5-9. For the migration of product data, the analytical approach architecture would be the most suitable. With analytical approach, the product data would be extracted from the source systems into a master data repository. From there, the data would be transformed and cleansed, and then loaded into the *new Product Register*.

After this, the Sinfos transition should be easy as the *new Product Register* could deliver the required XML documentation for Sinfos and it would include all the new product data values that were introduced along with the MR 6.0. Obviously, the *new Product Register* requires testing and user training before implementation. And during the WS1 to WS2 transition in Sinfos, the functionality of *new Product Register* and its output files for Sinfos should be monitored. Alike in the new Product Register and in the SAP, some history data is needed for the information systems renewals. The history data is relevant so that the ITF can carry out its business smoothly without any setbacks.

For the SAP implementation and Product Register renewal, the use of ITF's current ERP-interface should be emphasized. Currently, the ERP-interface provides certain product data for business supporting information systems, depending on the system. The use of the ERP-interface excludes the direct interaction between business supporting systems and the NAV/Product Register, all the data is connected through the ERP-interface. So, when renewing the Product Register, the *new Product Register* should be implemented in the same place as the old one, and configured in a way that the ERP-interface can provide the same services as usual. This is the ideal situation from interface perspective as the service execution is hidden from the service users.

The same idea of using the ERP-interface applies to SAP as well, though there are some information systems that interact with current ERP-system, i.e. NAV, directly, without the ERP-interface (*see appendix 2*). Therefore, the SAP implementation requires more effort as there is interaction with and without the ERP-interface. Though, it may come to the point where the whole ERP-interface is renewed and it changes the status quo. Also, it depends on the other information systems and their future, i.e. will they be replaced or removed in near the future. Known fact is that the MDM Kalido's master data will be used for data import to SAP. These are the relations and requirements that have to be considered when implementing SAP. Moreover, it has to be mapped that what product data will be received from Kalido and what has to be updated manually.

In conclusion, here are the compiled answers for research questions, presented in the chapter 1.2. First, answering to the four supporting research questions and finally answering for the main research question.

In which information systems the current product data is located?

Currently the product data is located in the business enabling information systems – NAV, Product Register, Kalido & Siebel – and in Sinfos. There is also several business supporting information systems, though their product data content is originated from the business enabling information systems. The appendices 5-9 encompass the locations of product data and product data contents per information system.

In which form the current product data is available in the information systems?

The availability of product data depends on the information system and users rights. For NAV, Product Register and Siebel there are no simple data export options for regular users, excluding the system administrator level. For Kalido, there are data export options as it can be seen from the appendix 8. Sinfos includes the largest variety of data export options, such as pdf-, excel- and txt-files. As some information systems lack of export functionality, the availability of product data is enabled by manual copying.

What kind of product data is needed for renewed information systems?

For the Product Register renewal, relevant product data should be migrated from the current business enabling information systems. Also the product data within Sinfos, especially the new requirements, should be included and notified. For the Sinfos renewal, the current Product Register and the SQL-queries encompass most of the product data imported for Sinfos. In addition, the new MR 6.0 requirements for Sinfos must be considered. Lastly, for the NAV renewal with SAP, the possibilities and product data content of MDM Kalido has to be studied, in order to map out the missing product data that has to be provided elsewhere, e.g. *new Product Register*.

How product data management should be arranged considering the information system renewals and MDM Kalido?

As the Product Register renewal will be the next major change, the PDM should be arranged around the *new Product Register* as it would be the major product data repository. By this, all the product data would be located there and it would serve as primary source for product data. The *new Product Register* can answer for the Sinfos requirements changes, if it is made for the Sinfos compatibility. The SAP implementation raises the question for MDM Kalido's purpose. The Kalido will be used for its product master data import to SAP. And for the SAP, it should be linked to ITF's information systems, especially to the *new Product Register*.

Finally, the main research question: *How company should arrange its product data management?*

The four supporting research questions already answered partly to this main question and final summary should be divided between PDM processes and information systems perspectives. From PDM process perspective, the processes for product data related functions should be revised and documented properly. This includes education of employees, enhancement of information sharing and segregation of duties. All this has to be done more transparently, so that each employee knows his or hers areas of responsibility and knows what other employees are responsible for. With this, overlapping of certain work functions and actions can be avoided, and in addition, employees will know the person who they have to consult in case of problems and challenges. Also, the perk of having a person in charge of PDM would be great, as that person would manage and guide the PDM functions and initiatives, instead of, having several employees who would be inventing the wheel separately.

From the information systems perspective, information systems should be used in order to enable and support the PDM functions and activities. With current information systems, the situation is manageable though inefficient due to the excessive manual work. As the Product Register will be renewed and NAV will be replaced with SAP, the information systems will provide better functionalities for product data and linkage between ITF and ITG. Along with information systems renewal and replacement, ITF should exploit the current ERP-interface ideology and configure it for the upcoming changes. Therefore the other information systems (business supporting information systems) behind the ERP-interface will be functioning as usual, though the execution of delivered services, by the ERP-interface, will be altered on the other side of the ERP-interface (*new Product Register* and SAP). The *new Product Register* as the primary product (master) data source for ITF, it will enhance employees' effectiveness as the product data availability and reliability increases. The SAP is linked with MDM Kalido, therefore minimizing some product data updates as the Kalido's product master data will be imported from Kalido to SAP. Also, the reporting for any product data related reports will be eased, as the ITF and ITG will be using same product numbers and identifiers (RSKU & SSKU).

7.2. Assessment of the study and further research

In this chapter the assessment of the study is presented along with the contributions and the discussion for further research. Chapter 7.2.1 presents the assessment of the study by going over the starting point of the research, covering research methods, philosophies, objectives, etc., and ending up with the assessment of the conclusions and recommendations. Chapter 7.2.2 presents the contributions of this research and further research topics and concepts.

7.2.1. Assessment of the study

The research began with the given research topics from the ITF. The preliminary research topics included several possibilities for distinct master's theses. Therefore, the topic had to be outlined and defined more carefully. Finally, the topic focused on product data, data migration and information systems renewal. The selection of these areas for this research was quite straight forwarded, as those issues were the most urgent for ITF. With this topic, research problem, questions, objectives and focus were selected with the help of supervising professor and the project sponsor from the ITF.

In addition to these research arrangements, research methods and philosophies were chosen. The research philosophy is based on interpretivism, which is ideal for understanding and interpretation of the research problem. To confirm the interpretivism as a research philosophy; combination of interpretive and functionalist paradigm were chosen, ontologically the research is subjective as researcher is part of the studied phenomena, epistemological choice is interpretivism and axiology lead to subjective view and interpretivism.

For research strategy, case study was chosen as the research focused on one case company, ITF. Lastly, the nature of research was chosen to be qualitative as it is suitable option together with interpretivism, and the purpose of the research was focused on descriptive and exploratory perspectives with features of explanatory perspective. As a researcher, it felt that the selection of research philosophy and other arrangements were quite difficult, as those concepts were quite unfamiliar. Though, it must be admitted that the thorough understanding and examination of above described concepts led to focused research and scientific evaluation of the results.

When compared the initial situation of research arrangements, methods, philosophies and strategies, to the final conclusions and recommendations for the ITF, it should be noted that the chosen research arrangements, methods, philosophies and strategies worked as intended for this research. The situation of research environment required interpretive approach and the encountered findings were analyzed with qualitative methods. These were obvious choices for the approach and results analyzing as the given research problem was qualitative in nature and therefore the quantitative methods or positivist philosophy could not fit for this research's purposes. Ontologically, subjective aspect was fitting, as the researcher worked in the researched environment. For purposes of the research, it described ITF's current situation, explored the possibilities for ITF's future measures and explained the relations with distinct entities.

In theoretical section, four topics were introduced: knowledge management, information systems, PDM and MDM. Knowledge management was included as the study covered the management of data and the enhancement of PDM practices within ITF. Information systems was included as the research related to the information systems renewal, and one of the sub topics was data migration, which was seen necessary component for comprehension of the research problem and for the given recommendations. PDM was included as a concept as the research examined the product data and its management within ITF. MDM was included in addition to the PDM, as the MDM gives slightly distinct perspective on data management.

Theoretical section provided sufficient knowledge for the research's empirical section, in order to answer for the research questions. Knowledge management revealed basics for knowledge creation and framework for its management. Information systems enlightened the process of IS renewals and implementations. For the PDM, theory was somehow ambiguous, as one spoke of PDM and another spoke about PLM, and some saw PDM as a part of PLM. Alas, the concept of PDM was emphasized for research's data perspective. MDM gave different and more detailed perspective to the data management than PDM. The MDM stressed the importance of business critical data, i.e. master data, for business exploitation, which was beneficial in conjunction with PDM.

For theoretical section, scientific literature of PDM and MDM were discovered as challenging topics due to their limited availability and their usability considering the settings of this research. For example, PDM is closely linked to the manufacturing functions, which were irrelevant for the research and its objectives. Despite mentioned challenge, scientific literature was gathered sufficiently in order to depict the concepts of PDM and MDM.

Empirical section and the data collection were constructed with interviews, consisting of two interviewing rounds. This served the purpose of data collection as it provided the best possible findings compared to other options. By this, is referred to the ease of data collection and the possibility of the researcher to ask additional questions during the interviews, as referred to the surveys where the options for answers has to be constructed in advance and there is only little room for the researcher to adapt within the data collection stage.

The interviews were semi-/structured theme interviews, which provided mostly qualitative data. Though, the qualitative data lead to division of interviews into two sections as the need for specific data was identified. This refers to the data content within each information system, and it was seen that the interviewees are more capable of giving more precise answers regarding to the data content when they can give their answers with time and proper preparation. Therefore, the data content per information system –questions were included within first section.

Interesting observation during the interviews was done, as it was noticed that the area of expertise for each interviewee affected their willingness to answer for certain questions within given theme, e.g. SAP-experts answers for PDM related issues, mostly consisting of opinion-based questions. As the focus is on certain area, it is difficult to give answers outside of your area of expertise. Though and strangely, for the ITF employees, this was not a similar issue. Still, the required answers were gotten when insisted and encouraged. For the usage of term PDM over MDM, it was seen more practical as the term of product master data could have distracted the interviewees as a complex term, whereas the product data and its management was more familiar term. Thus, avoiding the possible distractions, this encumbered researcher's work for data analysis and this was seen less troublesome than the distraction of the interviewees.

Qualitative analysis was used for data analysis, and the analysis was conducted according the occurring themes. By this, the data was analyzed and presented according certain themes, which lead to rational conclusions and recommendations. In retrospection, the interviews provided the needed data for the research, though the interviews could have been briefer (though, everyone can be wise after the event). Alas, it must be noticed that the performance of the interviews could have been smoother, if the interviews had been executed when scheduled and if the preparation for the first section could have been slightly better for some interviewees, e.g. forgotten interview, blank questionnaire, lost email. To be frank and honest, this was expected at some level and it was taken into consideration, though the avoidance of these issues could have required a lot of effort, so the risk was consciously taken. Overall, the execution of interviews was successful though there is some room for improvement due to occurred delays and misapprehensions.

For conclusions and recommendations, empirical section provided sufficient information and it was supplemented with the information of the theoretical section. As the research was conducted as a case study and the case company's situation was quite unique, it was difficult to find out model solutions from the theoretical literature. Thus, the given recommendations are based on the insights of theory and empirical evidence. For the ITF, the research provided the mapping of product data and its whereabouts; it depicted the used information systems and their relations, examined the product related functions for improvement and studied the upcoming changes for ITF and related parties. The recommendations along with this research give a comprehensive description of ITF's current status, exploration to the upcoming IS and product data alterations and explain the missing gap between current and desired state. Thus, the given recommendation is the best possible option for the time being and with these resources, as there is no universal or correct answer.

7.2.2. Contributions and further research

The greatest contribution of this research is essentially for the case company ITF. Based on this research and its results and recommendations, ITF possess a set of guidelines and directions for the upcoming changes, i.e. the renewal of Product Register with Sinfos linkage, and the implementation of SAP as ITF's new ERP-system. The most relevant contribution is the depiction and mapping of current situation and the compilation of future events that will change and alter ITF’s current situation. Furthermore, the research contributes recommendations for the missing gap between the current and future states.

For the theoretical perspective, research’s contribution for theoretical literature is minimal as the research was conducted as a case study, and therefore its findings and conclusion are not generally applicable for other instances. The application of these findings and conclusions would require an event that would be similar to the case company’s situation. However, smaller entities, such as arrangement of PDM with given starting point and limitations, can be applied for other instances, in certain extend of course.

Regardless of the minor contribution to the scientific and theoretical literature, there some topics and issues for further research. Distinction for further research, from the theoretical basis – IS renewals, data migration, PDM, MDM – and for the case company perspective, can be made. For the theoretical viewpoint, conjunction of PDM and MDM could be examined, as well as the question of IS renewals and rearrangement of business process related to the IS renewal. That is, should the IS and processes change simultaneously or separately, and if so, then which should be dealt first and which last. From the case company perspective, following topic could be studied: how the IS renewals and data migrations were executed, was there any challenges or success factors, how these issues reflect to the scientific literature? Also, the significance and effect of the PDM or MDM initiatives could be studied from efficiency perspective or from the productivity perspective.

BIBLIOGRAPHY

Ameri, F. & Dutta, D. 2005. Product Lifecycle Management: Closing the Knowledge Loops. Computer-Aided Design & Applications, vol. 2. no. 5. pp. 577-590

Ardagna, D., Francalani, C. & Piuri, V. 2004. Designing and Rightsizing the Information System Architecture. Information Systems Frontiers, vol. 6, no. 3, pp. 229-245

Berghout, E., Nijland, M. & Powell, P. Management of lifecycle costs and benefits: Lessons from information systems practice. Computers in Industry, vol. 62, no. 7, pp. 755-764

Breakfield, C.V. & Burkey, R.E. 2002. Chapter 5 – The Migration Process in Managing Systems Migrations and Upgrades. Woburn, Digital Press, pp. 145-201.

Cantamessa, M., Montagna, F. & Neirotti, P. 2012. Understanding the organizational impact of PLM systems: evidence from an aerospace company. International Journal of Operations & Production Management, vol. 32, no. 2, pp. 191-215

Carson, D., Gilmore, A., Perry, C. & Grønhaug, K. 2001. Qualitative Marketing Research: Philosophy of Research. Sage Research Methods. 16 p.

Celko, J. 2006. Joe Celko's Analytics and OLAP in SQL. Burlington, Morgan Kaufmann. 208 p.

CIMdata, 1997. Product data management: The definition – An Introduction to concepts, beliefs, and terminology. [WWW] [Accessed 12.7.2012] Available: http://www.google.fi/url?sa=t&rct=j&q=cimdata%20product%20data%20definition&source=web&cd=3&sqi=2&ved=0CFUQFjAC&url=http%3A%2F%2Fcypress.kaist.ac.kr%2Fres%2F2011%2520IE553%2FPDMDef.pdf&ei=wHv-T-C_N_PE4gTP0PT8Bg&usg=AFQjCNF7G7LM9WiH32vpO73WBxpDrtxmhA&cad=rja. 27 p.

CIMdata. 2002. Product lifecycle management: Empowering the future business. [WWW] [Accessed 19.11.2012] Available: http://doc.mbalib.com/view/8cb8b763acd25742ed7e72ad9ac29d01.html

Collins English Dictionary. 2013. MOT Dictionaries. [WWW, restricted access] [Accessed 23.01.2013] Available: http://mot.kielikone.fi/mot/ttkk/netmot.exe?motportal=80

Do, N. & Chae, G. 2011, A Product Data Management architecture for integrating hardware and software development. Computers in Industry, vol. 62, no. 8–9, pp. 854-863.

Eriksson & Kovalainen 2008. Qualitative methods in business research – research philosophy. Sage Research Methods. 13 p.

Eynard, B., Gallet, T., Roucoules, L. & Ducellier, G. 2006. PDM system implementation based on UML. Mathematics and Computers in Simulation, vol. 70, pp. 330-342.

Ghauri, P. & Grønhaug, K. 2005. Research Methods in Business Studies – A Practical Guide. 3rd edition. Harlow, Prentice Hall. 257 p.

Gimenez, D.M., Vegetti, M., Leone, H.P. & Henning, G.P. 2008. PRoduct ONTOlogy: defining product-related concepts for logistics planning activities. Computers in Industry, vol. 59, pp. 213-241.

GS1. Extranet, development notifications & e-mail conversations. [WWW, restricted access] [Accessed 26.02.2013] Available: http://www.gs1.fi/extranet

Haikala, I. & Märijärvi, J. 2006. Ohjelmistotuotanto. 11th edition. Helsinki, Talentum, 440 p.

Irani, Z., Themistocleous, M. & Love, P. E. D. 2003. The impact of enterprise application integration on information system lifecycles. Information & Management, vol. 41, no. 2, pp. 177-187

Jonkers, H., Lankhorst, M. M., ter Doest, H. W. L., Arbab, F., Bosma, H. & Wieringa, R. J. 2006. Enterprise architecture: Management tool and blueprint for the organization. Information Systems Frontiers, vol. 8, no. 2, pp. 63-66.

Kamara, J. M., Anumba, C. J. & Carrillo, P. M. 2002. A CLEVER approach to selecting a knowledge management strategy. International Journal of Project Management, vol. 20, no. 3, pp. 205-211

Kelly, C. & Nelms, C. 2003. Roadmap to checking data migration. Computers & Security, vol. 22, no. 6, pp. 506-510.

Kropsu-Vehkapera, H., Haapasalo, H., Harkonen, J. & Silvola, R. 2009, Product data management practices in high-tech companies. Industrial Management & Data Systems, vol. 109, no. 6, pp. 758-774.

Kurbel, K. E. 2008. The Making of Information Systems. Chapter 3 – Information Systems Architecture. Springer Berlin Heidelberg, pp. 95-154.

Liebowitz, J. & Megbolugbe, I. 2003. A set of frameworks to aid the project manager in conceptualizing and implementing knowledge management initiatives. International Journal of Project Management, vol. 21, no. 3, pp. 189-198.

Liu, D.T. & Xu, X.W. 2001, A review of web-based product data management systems. Computers in Industry, vol. 44, no. 3, pp. 251-262.

Loshin, D. 2009. Master Data Management. Morgan Kaufmann, Boston, 259 p.

Loshin, D. 2011. The Practitioners Guide to Data Quality Improvement. Morgan Kaufmann, Boston, 383 p.

Nonaka, I. & Takeuchi, H. 1995. The Knowledge-Creating Company – How Japanese Companies Create the Dynamics of Innovation. New York, Oxford University Press. 284 p.

Otto, B. 2012. How to design the master data architecture: Findings from a case study at Bosch. International Journal of Information Management, vol. 32, no. 4, pp. 337-346.

Oxford Reference Online. 2013. [WWW, restricted access] [Accessed 23.01.2013] Available: http://www.oxfordreference.com/view/10.1093/oi/authority.20110803095422245

Premier International. 2004. Rapid Application Development (RAD) for Data Migration – White Paper Solutions by Premier International. [WWW] [Accessed 28.01.2013] Available: http://www.premier-international.com/pdf/Applaud_White_Paper.pdf

Saaksvuori, A. & Immonen, A. 2008. Product Lifecycle Management. [WWW] [Accessed 19.11.2012] Available: http://link.springer.com/book/10.1007/978-3-540-78172-1/page/1

Saunders, M., Lewis, P. & Thornhill, A. 2009. Research methods for business students. 5th edition. Harlow, Prentice Hall. 614 p.

Sinfos. Sinfos Product Data Bank. [WWW, restricted access] [Accessed 04.03.2013] Available: http://portal.sinfos.de/cataloom/sinfosportal/index.do

Sydänmaanlakka, P. 2004. Älykäs organisaatio – Tiedon, osaamisen ja suorituksen johtaminen. 7th edition. Helsinki, Talentum. 303 p.
University of Jyväskylä. 2012. Tieteenfilosofiset suuntaukset. [WWW]. [Accessed 09.10.2012]. Available: https://koppa.jyu.fi/avoimet/hum/menetelmapolkuja/menetelmapolku/tieteenfilosofiset-suuntaukset
Thalheim, B. & Wang, Q. 2013. Data migration: A theoretical perspective. Data & Knowledge Engineering, [WWW, Available online 23.12.2012] [Accessed 06.01.2013] Available: http://dx.doi.org/10.1016/j.datak.2012.12.003.
Thierauf, R. J. 2001. Effective business intelligence systems. [WWW, restricted access] [Accessed 05.01.2013] Available: http://site.ebrary.com/lib/ttyk/docDetail.action?docID=10005668
Turban, E., Leidner, D., McLean, E. & Wetherbe, J. 2008. Information technology for management: transforming organizations in the digital economy. 6th edition. Hoboken (N.J.), Wiley. 662 p.
von Wright, G. H. 1970. Tieteenfilosofian kaksi perinnettä. Reports from the Institute of Philosophy. No. 1. University of Helsinki. 32 p.

APPENDIX 1A: QUESTIONNAIRE FOR SECTION I – CATEGORY I (1/2)

Background

1. Who are you?
2. Educational background (technically oriented, etc.)?
3. What is your position/title in this company?
4. What do you do (job description)?
5. Which applications do you use that uses product data? And are you a basic user or a main user (responsible for update actions)?

Applications – answer for every application (product data perspective), that you use, separately.

1. What is application X's main function?
2. How the product data is imported/updated to application X? Is it done automatically, manually or both and?
3. Where the imported/updated product data is gathered/received?
4. How the updated product data is gathered/received?
5. How often application X's product data has to be updated?
6. If two or more information systems contain the same product data, in case of update, will the information be updated automatically to all relevant information systems?
7. If the update is not done automatically, then how the update is done to all relevant information systems?

Existing product data – answer for every application (product data perspective), that you use, separately.

1. What product data is in application X?
2. Are there available documents on the existing product data?
3. If not, how do we know that what product data is in application X?
4. Does application X include all the required product data for ITF's business operations? From the application's perspective: data usage, update actions and data distribution.
5. If it does not, then where the supplementary product data is gathered?

APPENDIX 1A: QUESTIONNAIRE FOR SECTION I – CATEGORY I (2/2)

Product data needs and requirements – answer for every application (product data perspective), that you use, separately.

1. How the product data requirements for application X are determined (from business perspective, i.e. what is needed to enable business)?
2. If there are any documents, where they could be found?
3. If there are not documents, then where the guidelines and norms of the product data requirements come from?
4. How the product data requirements for application X are determined (from information system perspective, technical viewpoint)?
5. If there are any documents, where they could be found?
6. If there are not documents, then where the guidelines and norms of the product data requirements come from?

PDM – Product Data Management

1. What is your opinion about PDM? How do you see it?
2. From your point of view, what is PDM's objective in a business context, i.e. in company's short and long term operations?
3. What are the key points for successful PDM?
4. And what are the major challenges with PDM?
5. How is the PDM arranged for ITF?
6. How the ITG and ITF's PDM should be linked together?
7. How often you are working with product data?
8. Do you update the product data or are you simply using it?
9. From your point of view, how would you describe/depict the current situation with the product data (its whereabouts, quality, availability, reliability, validity)?
10. How do think/see/know that how the product data is organized and managed in ITF?
11. And how it (PDM) should be organized and managed (from your perspective)?
12. How the ITG and its policies effect on PDM?
13. Are there any limitations or requirements that ITG sets to be followed for product data?

APPENDIX 1B: QUESTIONNAIRE FOR SECTION I – CATEGORY II (1/2)

Background

1. Who are you?
2. Educational background (technically oriented, etc.)?
3. What is your position/title in this company?
4. What do you do (job description)?
5. Which applications do you use that uses product data?

General issues about product data (IS perspective)

1. In which applications product data is used?
2. Which applications and IS functions are hosted internally?
3. Which applications and IS functions are hosted externally?
4. Are there centralized actions for product data updates?
5. In case of product data update (insertion – new product, correction, deletion), how and where the update actions are focused?
6. In case of product data update, where update information comes from? And does the information content and source differ depending on the update actions?
7. What requirements applications set for product data?
8. Does the product data requirements differ between IS and business perspectives?

APPENDIX 1B: QUESTIONNAIRE FOR SECTION I – CATEGORY II (2/2)

PDM – Product Data Management

1. What is your opinion about PDM? How do you see it?
2. From your point of view, what is PDM's objective in a business context, i.e. in company's short and long term operations?
3. What are the key points for successful PDM?
4. And what are the major challenges with PDM?
5. How is the PDM arranged for ITF?
6. And how the ITG and its ITF's PDM should be linked together?
7. How often you are working with product data?
8. Do you update the product data or are you simply using it?
9. From your point of view, how would you describe/depict the current situation with the product data (its whereabouts, quality, availability, reliability, validity)?
10. How do think/see/know that how the product data is organized and managed in ITF?
11. And how it (PDM) should be organized and managed (from your perspective)?
12. How the ITG and its policies effect on PDM?
13. Are there any limitations or requirements that ITG sets to be followed for product data?
14. At the moment ITF does not operate on SAP as their ERP-system. How the MDM Kalido's product data is linked to ITF's current information systems?

APPENDIX 1C: QUESTIONNAIRE FOR SECTION I – CATEGORY III (1/2)

Background

1. Who are you?
2. Educational background (technically oriented, etc.)?
3. What is your position/title in this company?
4. What do you do (job description)?

PDM – at a general level

1. What is your opinion about PDM? How do you see it?
2. From your point of view, what is PDM's objective in a business context, i.e. in company's short and long term operations?
3. What are the key points for successful PDM?
4. And what are the major challenges with PDM?
5. How is the PDM arranged for ITG? Any complications, future projects, etc?
6. And how the PDM should be arranged in ITG's subsidiaries?
7. And how the ITG and its subsidiary's PDM should be linked together?

MDM Kalido

1. MDM Kalido is for corporate master data management, how it is related to ITG's subsidiaries?
2. And how about the product data perspective, how it affects ITG's subsidiaries? Does the updated product data update automatically to the subsidiaries' information systems?
3. What product data is stored in MDM Kalido?
4. In which form the product data in Kalido is exportable?
5. Is any other subsidiary of ITG using SA2? And if they are, then are they using MDM Kalido for product data import?
6. At the moment ITF does not operate on SAP as their ERP-system. How the MDM Kalido's product data is linked to ITF's information systems?
7. Is the product data content in MDM Kalido same for every subsidiary of ITG?
8. If there are any variations, then why it is so and will it be possible to modify ITF's product data content in MDM Kalido?

APPENDIX 1C: QUESTIONNAIRE FOR SECTION I – CATEGORY III (2/2)

SAP & Product data

1. Speaking of ERP (Enterprise Resource Planning) systems on general level, how well they adapt to product data storing and distribution for other applications?
2. Is ERP system enough for company's product data upkeep or are other solutions necessary?
3. ITG has lined that SAP-ERP must be used in its subsidiaries. Is SAP and MDM Kalido compatible with each other (for product data exchange)?
4. If SAP and Kalido are compatible, then are the interfaces and product data contents same for every subsidiary?

APPENDIX 1D: QUESTIONNAIRE FOR SECTION II – CATEGORY I

General – answer for every application (product data perspective)

1. How the product data should be arranged/organized for you to work more efficiently? What would you change and why?
2. Are you facing any problems with the product data in your daily routines?
3. Are you aware of actions/updates that will affect product data's requirements and its management and content?
4. There are different SSKU's (*Standard Stock Keeping Unit*) and RSKU's (*Retail Stock Keeping Unit*) for ITG and ITF. How that affects product data and its management?

Product data

1. Do you use any supplementary documents in your daily work with product data, i.e. Excel-spreadsheets or similar?
2. If you use, then who have created those documents and who is responsible for their upkeep?
3. Is there something that has been done well with the product data?
4. Which data sources do you use for reporting (reporting related to product data)?

ERP – SAP

1. What do you expect from the SAP, from product data perspective? Any changes from Navision to SAP? Any wishes?

Sinfos-interface + Product Register (replacement)

1. The Sinfos interface is going to change, so there is no more TXT-interface after 2/2014 (ITF uses TXT-interface). The Product Register makes input files for the TXT-interface, for the Sinfos import. And the SAP-project is planned to start 10/2013 and to end 10/2014 (product data has to be stored somewhere else than SAP in that time period)? What solution you see as a possible substitute for Product Register i.e. what should be done?

Theoretical model vs. practical model

1. Here is the theoretical model for PDM. How does it fit for ITF's framework considering the given limitations and presence of ITG?
2. Which/What changes or modifications you would do to this model, so it would fit for ITF?
3. Now you have created working solution for ITF's PDM, will it work in the future (in case of changes, like SAP)?

APPENDIX 1E: QUESTIONNAIRE FOR SECTION II – CATEGORY II (1/2)

ERP – SAP

1. What product data can be stored in SAP?
2. Is it possible to store all product data in SAP?
3. If possible, will it be plausible (too expensive perhaps)?
4. If not, then what supplementary solution will be acquired in addition to SAP?
5. SAP project is planned to start October 2013; does it make any restrictions for product data?
6. What are the import/export options for product data in SAP?
7. Probably there is going to be a SAP-interface for product data retrieval from SAP?
8. If so, is it easy to modify for different and changing needs?
9. What differences exist between Navision and SAP?
10. What are the benefits of SAP implementation for ITF?
11. How product data updates and insertions will change after SAP implementation (difference between NAV and SAP)?
12. How easily SAP can be modified for possible product data requirement changes?
13. For SAP implementation, from where the product data will be imported?

APPENDIX 1E: QUESTIONNAIRE FOR SECTION II – CATEGORY II (2/2)

Sinfos-interface + Product Register (replacement)

1. Sinfos interface is going to change, which options you see as a possible solutions for new interface (XML | Excel | Pricat (EDI))?
2. The Sinfos interface is going to change, so there is no more TXT-interface after 2/2014 (ITF uses TXT-interface). The Product Register makes input files for the TXT-interface, for the Sinfos import. And the SAP-project is planned to start 10/2013 and to end 10/2014 (product data has to be stored somewhere else than SAP in that time period)? What solution you see as a possible substitute for Product Register i.e. what should be done?
3. Are there any applications or IS functions that will be hosted internally instead of externally, and vice versa?
4. There are different SSKU's (*Standard Stock Keeping Unit*) and RSKU's (*Retail Stock Keeping Unit*) for ITG and ITF. How that affects product data and its management?

Theoretical model vs. practical model

1. Here is the theoretical model for PDM. How does it fit for ITF's framework considering the given limitations from ITG?
2. Which/What changes or modifications you would do to this model, so it would fit for ITF?
3. Now you have created working solution for ITF's PDM, will it work in the future (in case of changes, like SAP)?

APPENDIX 1F: QUESTIONNAIRE FOR SECTION II – CATEGORY III (1/2)

ERP – SAP

1. What product data can be stored in SAP?
2. Is it possible to store all product data in SAP?
3. If possible, will it be plausible (too expensive perhaps)?
4. If not, then what supplementary solution will be acquired in addition to SAP?
5. SAP project will not start until October 2013; does it make any restrictions for product data?
6. What are the import/export options for product data in SAP?
7. Probably there is going to be a SAP-interface for product data retrieval from SAP?
8. If so, is it easy to modify for different and changing needs?
9. What differences exist between Navision and SAP?
10. What are the benefits of SAP implementation for ITF?
11. How the SAP and MDM Kalido are linked together? Are there standard interfaces for data exchange?
12. What is the functional benefit from SAP and MDM Kalido linkage, from business perspective?
13. How product data updates and insertions will change after SAP implementation (difference between NAV and SAP)?
14. How easily SAP can be modified for possible product data requirement changes?
15. For SAP implementation, from where the product data will be imported?
16. There has been a SAP implementation on Norway's office, is it possible to have case material about the implementation? Any success and/or critical factors?

APPENDIX 1F: QUESTIONNAIRE FOR SECTION II – CATEGORY III (2/2)

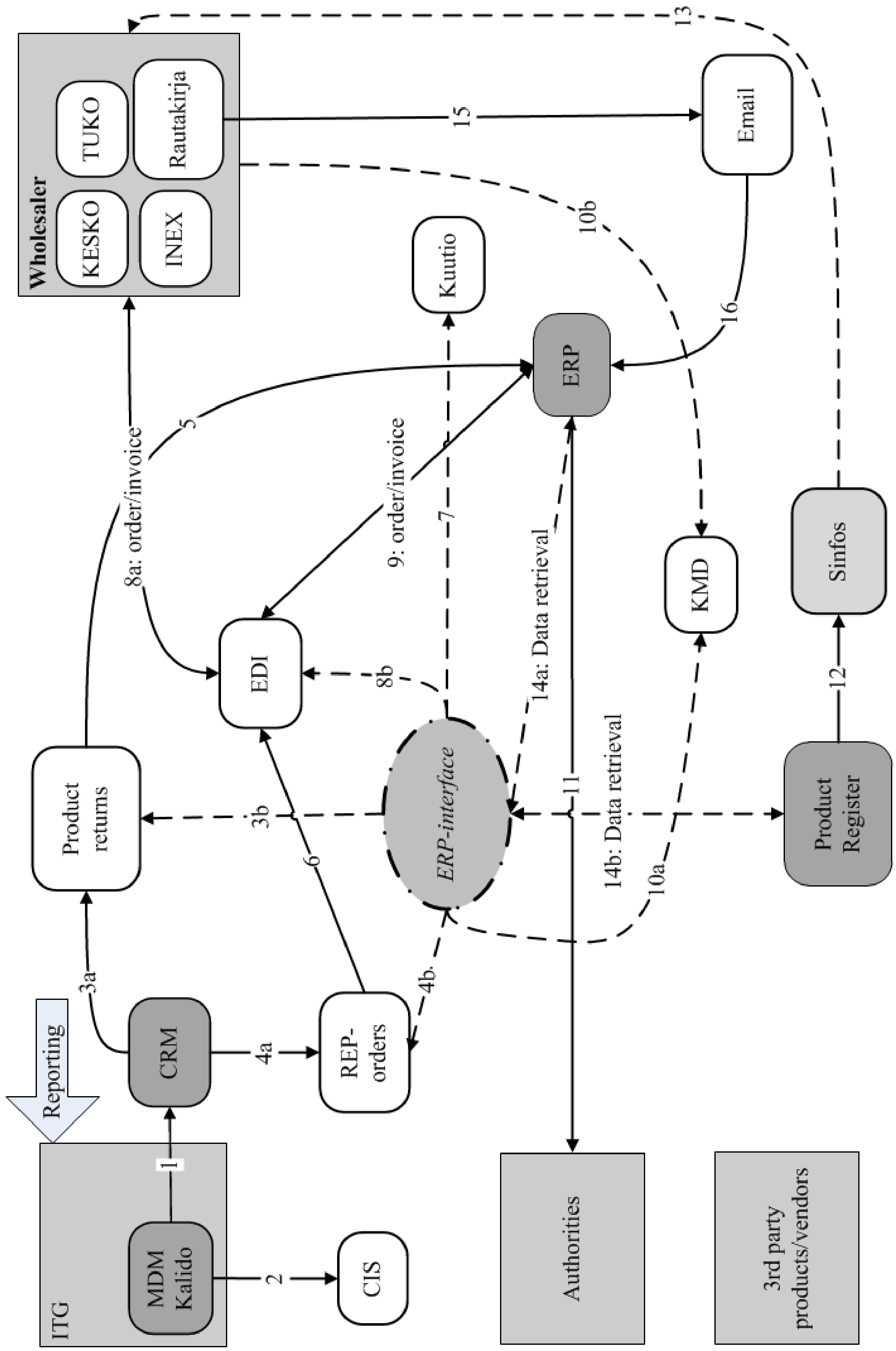
Sinfos + Product Register (replacement)

1. Considering the SAP project, it is planned to start in autumn 2013 but ITF needs working PDM system at the same time. So is there any solution that would be reasonable for PDM and for upcoming SAP implementation?
2. The Sinfos interface is going to change, so there is no more TXT-interface after 2/2014 (ITF uses TXT-interface). The Product Register makes input files for the TXT-interface, for the Sinfos import. And the SAP-project is planned to start 10/2013 and to end 10/2014 (product data has to be stored somewhere else than SAP in that time period)? What solution you see as a possible substitute for Product Register i.e. what should be done?
3. There are different SSKU's (Standard Stock Keeping Unit) and RSKU's (Retail Stock Keeping Unit) for ITG and ITF. How that affects product data and its management?
4. Should the ITF get rid of the local SKU's?
5. In case of using third party applications or services, what are the ITG restrictions and qualifications for this kind of settlement?
6. Are you aware of any ITG's subsidiaries which use SA2 (product data bank for wholesalers)?
7. If you are, do you know that how they are importing their product data to SA2?
8. Could MDM Kalido be used for this kind of product data import? i.e. is there enough data in Kalido?

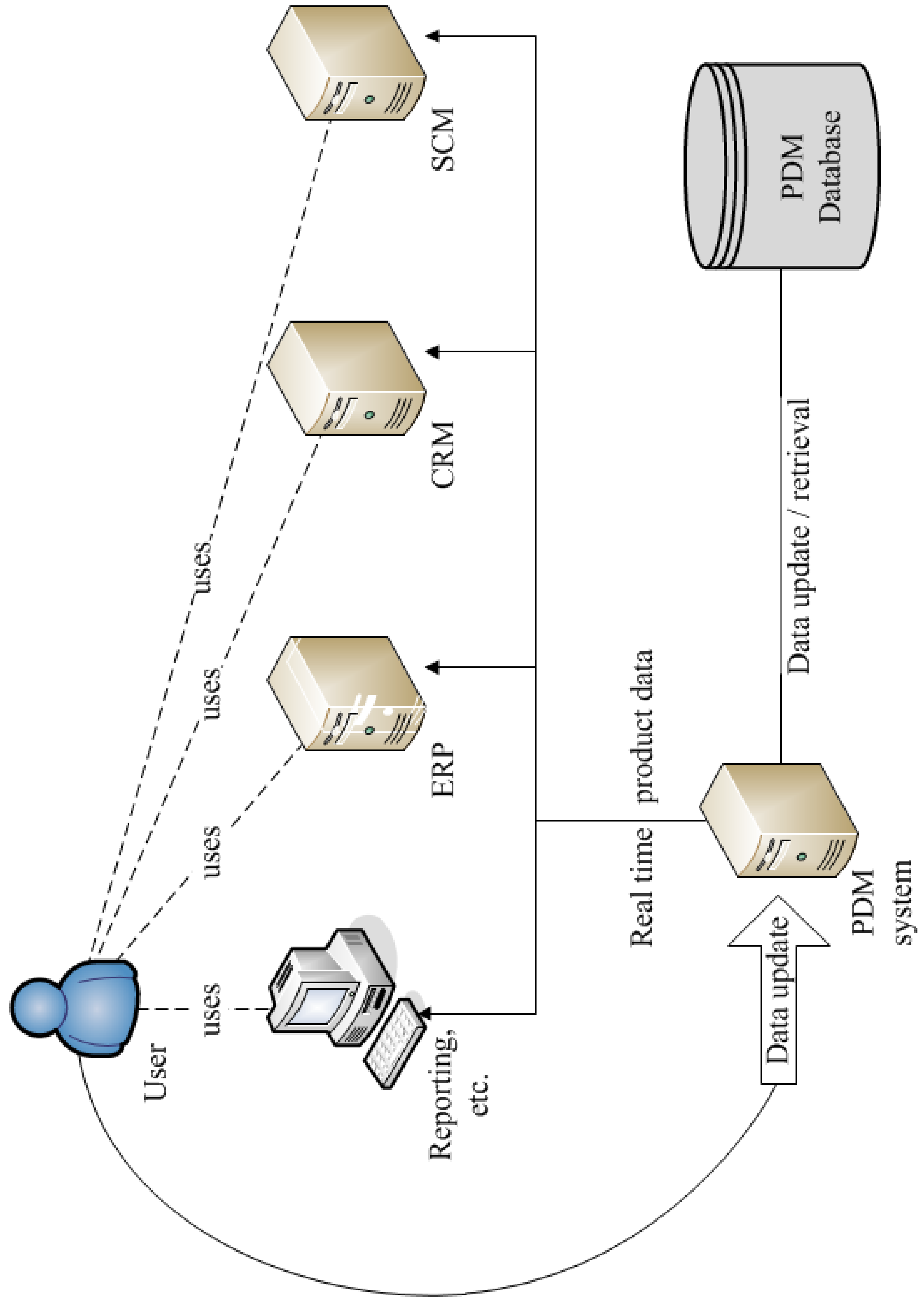
Theoretical model vs. practical model

1. Here is the theoretical model for PDM. How does it fit for ITF's framework considering the given limitations from ITG?
2. Which/What changes or modifications you would do to this model, so it would fit for ITF?
3. Now you have created working solution for ITF's PDM, will it work in the future (in case of changes, like SAP)?

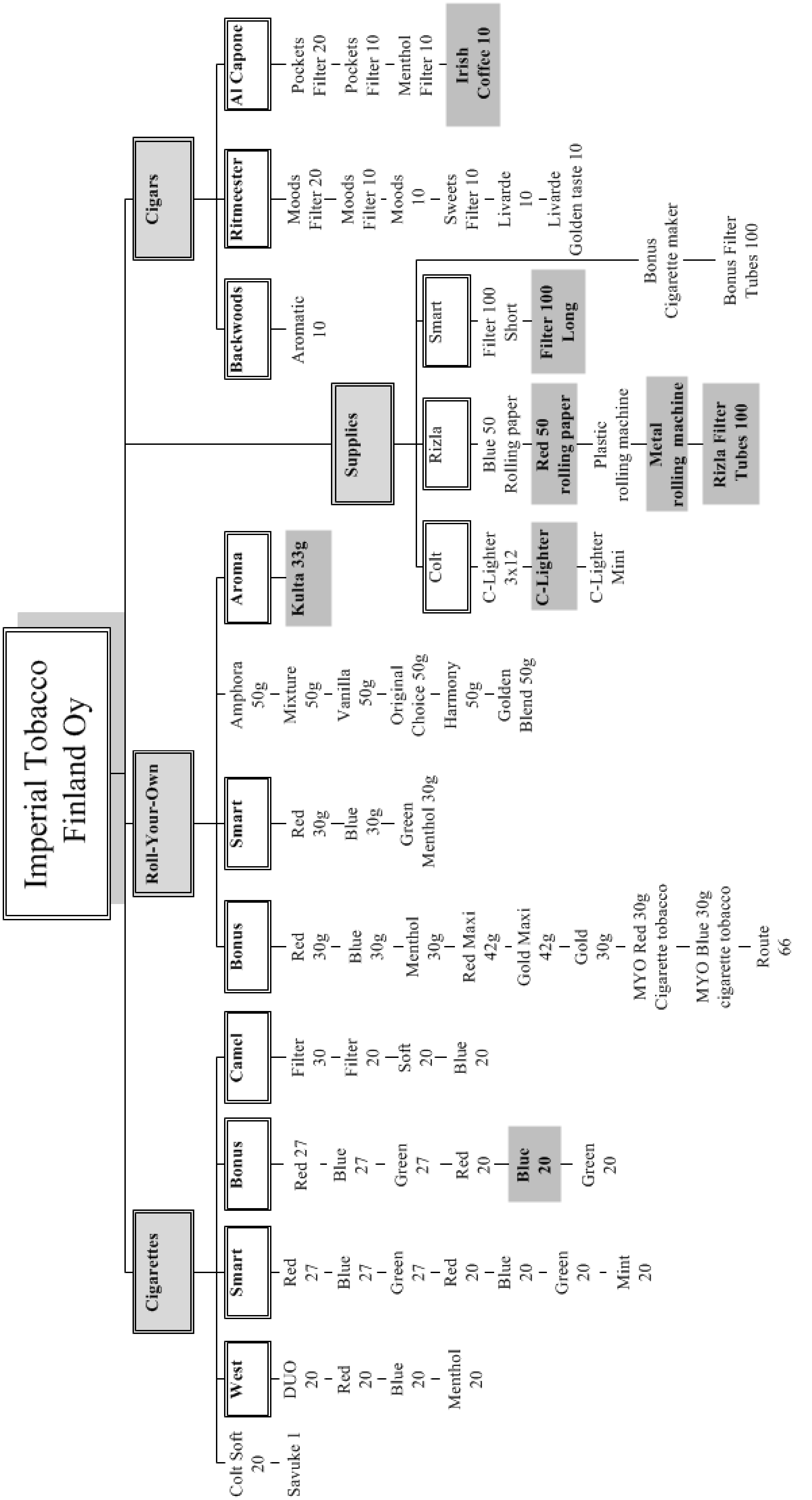
APPENDIX 2: DATA ARCHITECTURE



APPENDIX 3: THEORETICAL PDM-MODEL



APPENDIX 4: SELECTION OF GOODS



APPENDIX 5: PRODUCT DATA CONTENT IN MS DYNAMICS NAV (1/2)

MS Dynamics NAV	Possible duplicates			
Data field (FI)	Product Register	Siebel	Kalido	Sinfos
Nro	X		X	
Yleinen nimikenro	X			X
Kuvaus	X		X	
Kuvaus 2				
Perusmittayksikkö				
Nimikekategorian koodi	X			
Tuoteryhmäkoodi	X			
Toimittajan nro				
Toimittajan nimikenro		X	X	
Oston mittayksikkö				
Hakunimi				
Varasto				
Määrä ostotilauksessa				
Määrä myyntitilauksessa				
Tavaranimike				
Alkuperämaan/-alueen koodi	X			
Arvostusmenetelmä				
Yksikkökustannus				
Yleiskustannus (arvo)				
Välillinen kustannus-%				
Viimeinen välitön kusta...				
Standardi koodi (SSKU)		X	X	
Tuotteen yksilöivä koodi (RSKU)		X	X	
Yleinen tuotteen kirjau...				
Tuotteen ALV-kirjaus				
Varaston kirjausryhmä				
Laskutettu nettomäärä				
Salli laskualennus				
Myynnin mittayksikkö				
Aktiivisuus				
Kaasun määrä				
Bruttopaino				
Nettopaino				
Valm. Verotuotetyyppi				
Sijaintikoodi				

APPENDIX 5: PRODUCT DATA CONTENT IN MS DYNAMICS NAV (2/2)

Data field (FI)	Product Register	Siebel	Kalido	Sinfos
KRS: määrä mittayksikköä kohden				
KULP: määrä mittayksikköä kohden	X		X	
LAVA: määrä mittayksikköä kohden	X			
LTK: määrä mittayksikköä kohden	X		X	
ME: määrä mittayksikköä kohden	X		X	
MLE: määrä mittayksikköä kohden				
KRS: kuutiotilavuus				
KULP: kuutiotilavuus				
LAVA: kuutiotilavuus	X			
KRS: kuutiotilavuus				
LTK: kuutiotilavuus	X			
ME: kuutiotilavuus				
MLE: kuutiotilavuus				
KRS: paino				
KULP: paino	X			
LAVA: paino	X			
LTK: paino	X			
ME: paino	X			
MLE: paino	X			
Viittauksen tyypin nro				
Viittauksen nro	X			
Mittayksikön kuvaus				

APPENDIX 6: PRODUCT DATA CONTENT IN PRODUCT REGISTER (1/2)

	Product Register	Possible duplicates			
	Data field	NAV	Siebel	Kalido	Sinfos
Perustiedot - Perustiedot	Tuotteen nimi	X			
	Tuotenumero	X			X
	Nimen lyhenne			X	
	Yritys			X	X
	Tuoteryhmä	X			
	Markkina-alue				
	Vahvuus				
	Tuoteperhe				
	Hintaryhmä				
	Valmistaja				
	Tavararyhmä				
	Tuotealaryhmä	X			
	Segmentti				
Perustiedot - Hinnastotiedot	Voimaan pvm				
	Kulutuspakkaus			X	X
	ME			X	X
	APL			X	X
	Lava				
	SCALA tuotenumero	X		X	
	OVH FIM				
	OVH / EUR	X			
	Tekstikenttä				
	Yksikköä / kul.pakk	X		X	
	Kul.pakk / ME				
	ME /APL				
	APL / Lavakerros				
	Kerroksia / Lava				
Perustiedot - Hinnastotiedot	KPL / ME	X		X	
	KPL / APL	X		X	
	APL / Lava				
	MLE / Lava	X			
	ME / Lava				
	Hinnasto ID				
Sinfos perustiedot	Tuotenimi_Pitka	X			
	Yksikko				
	ValmisteveroLuokka				
	Tuotenimi_Brand_FI			X	X
	Tuotenimi_Brand_SW			X	X
	Tuotenimi_Brand_EN			X	X
	Tuotenimi_taydellinen_FI				
	Tuotenimi_taydellinen_SW				
	Tuotenimi_taydellinen_EN				
	CN koodi	X			
	Tavararyhmä				
	CPV				
	Alkuperämaa	X			
	Viimeisin valmistusmaa				

APPENDIX 6: PRODUCT DATA CONTENT IN PRODUCT REGISTER (2/2)

	Data field	NAV	Siebel	Kalido	Sinfos
Sinfos pakkaustiedot	Pakkaustyyppi_Kulpakk				
	Pakkaustyyppi_Me				
	Pakkaustyyppi_Apl				
	LavojaPaallekkain				
	SuojaAikaPaattyy (pp.kk.vvvv)				
Sinfos mitat	Piikkireika_koodi				
	Piikkireika_vaakaetaisyys				
	Piikkireika_pystyetaisyys				
Mitat ja painot	Kuljetuspakkaus: leveys				
	Kuljetuspakkaus: pituus				
	Kuljetuspakkaus: korkeus				
	Myyntierä: leveys				
	Myyntierä: pituus				
	Myyntierä: korkeus				
	APL: leveys				
	APL: pituus				
	APL: korkeus				
	Koko lava: leveys				
	Koko lava: pituus				
	Koko lava: korkeus				
	APL: m3	X			
	Koko lava: m3	X			
	Kuluttajapakkaus: netto	X			
	Kuluttajapakkaus: brutto				
	Myyntierä: netto	X			
	Myyntierä: brutto				
	APL: netto	X			
	APL: brutto				
	Koko lava: netto	X			
	Koko lava: brutto				
	kg /MLE	X			
	Lavan tyyppi				
Savuke	Pituus / mm				
	Stickin paino / g				
	Halkaisija / mm				
	Filtterin pituus / mm				
	Tippi pituus /mm				
	Terva / mg				X
	Nikotiini /mg				X
	CO /mg				X
	Tuotenumero: Meira Nova	X			
	Tuotenumero: Tuko	X			

APPENDIX 7: PRODUCT DATA CONTENT IN SIEBEL CRM

Siebel	Possible duplicates			
Data field	NAV	Product Register	Kalido	Sinfos
Product name	X		X	
Description				
Alias name			X	
Audit sequence				
Product category			X	X
Product level				
Parent product	X		X	
Organization		X	X	X
Product consumer price	X	X		X
Transfer order (checkbox)				
Auditable (checkbox)				
Active (checkbox)				
Sales product (checkbox)				

APPENDIX 8: PRODUCT DATA CONTENT IN MDM KALIDO (1/3)

MDM Kalido		Possible duplicates			
Data field	Example value	NAV	Product Register	Siebel	Sinfos
RETAIL_SKU_CODE	2102832	X		X	
RETAIL_SKU_NAME	Smart KS Blue Smooth LIP - Box 27 pcs Stick (6.30 ND HW) : FI			X	
RETAIL_SKU_S_NAME	Smart Blue LIP KS Box 27 (6.30 ND HW) FI				
RET_SKU_STOCK_CASE_CONT	5400.000000000000000	X	X		
RET_SKU_OUTER_CONT	216.000000000000000	X	X		
RET_SKU_VOLUME_CASE	44511.000000000000000				
STOCK_CASE_BARCODE	6410102698175		X		X
RET_PACK_BARCODE	6410102298115		X		X
OUTER_BARCODE	6410102498126		X		X
RET_SKU_NET_WEIGHT_CASE	5832.000000000000000				
RET_SKU_GROSS_WEIGHT_CASE	6523.000000000000000				
PROD_PLAN_STRAT_CODE	1				
PROD_PLAN_STRAT_NAME	MTS				
PROD_PLAN_STRAT_S_NAME	MTS				
RET_SKU_FCAST_TYPE_CODE	0				
RET_SKU_FCAST_TYPE_NAME	Standard Priced				
RET_SKU_STATUS_CODE	1				
RET_SKU_STATUS_NAME	Active				
STANDARD_SKU_CODE	029341	X		X	
STANDARD_SKU_NAME	Smart KS Blue Smooth LIP - Box 27 pcs Stick : FI				
STANDARD_SKU_S_NAME	Smart Blue Smooth LIP KS Box 27 FI				
MARKET_CONSTRAINT_CODE	043				
MARKET_CONSTRAINT_NAME	Finland				
MARKET_CONSTRAINT_S_NAME	FI				
MANUFACTURING_COMP_CODE	001				
PRODUCT_TYPE_CODE	1				
PRODUCT_TYPE_NAME	ITG Brand				
TRADEMARK_OWNER_CODE	001				
BRAND_PACK_VAR_CODE	020406				
BRAND_PACK_VAR_NAME	Smart KS Blue Smooth - King Size Lights no Filter Type American Blend Standard KS no Flavour Type Box 27 pcs Stick				
BRAND_PACK_VAR_S_NAME	Smart Blue Smooth KS Box 27 Stick				
RET_BNDL_SHPE_TYPE_CODE	02				
RET_PACK_CONT_TYPE_CODE	130				
CONSUMER_UNIT_CONTENT	27.000000000000000	X	X		X
RET_PACK_CONTENT	27.000000000000000	X	X		X
RET_PACK_TYPE_CODE	02				

APPENDIX 8: PRODUCT DATA CONTENT IN MDM KALIDO (2/3)

Data field	Example value	NAV	Product Register	Siebel	Sinfos
BRAND_VAR_CODE	007001				
BRAND_VAR_NAME	Smart KS Blue Smooth - King Size Lights no Filter Type American Blend Standard KS no Flavour Type				
BRAND_VAR_S_NAME	Smart Blue Smooth KS				
RET_LENGTH_BAND_CODE	03				
RET_LENGTH_BAND_NAME	King Size (83 - 87mm)				
RET_TASTE_TYPE_CODE	03				
RET_TASTE_TYPE_NAME	Lights				
RET_FILTER_TYPE_CODE	12				
RET_FILTER_TYPE_NAME	no Filter Type (Standard Filter)				
RET_DIAMETER_TYPE_CODE	03				
RET_DIAMETER_TYPE_NAME	Standard KS				
RET_BLEND_TYPE_CODE	02				
RET_BLEND_TYPE_NAME	American Blend				
RET_FLAVOUR_TYPE_CODE	12				
RET_FLAVOUR_TYPE_NAME	no Flavour Type				
BRAND_SUB_FAM_CODE	9013				
BRAND_SUB_FAM_NAME	Smart Cigarettes				
BRAND_SUB_FAM_S_NAME	Smart CIG				
CIGAR_PROD_SEGMENT_CODE	0007				
BRAND_FAM_CODE	000233				
BRAND_FAM_NAME	Smart Cigarettes				
BRAND_FAM_S_NAME	Smart CIG				
REP_BRAND_FAM_CODE	0096				
STRATEGIC_GROUP_CODE	03				
PRODUCT_GROUP_CODE	01				
PRODUCT_GROUP_NAME	Cigarettes			X	X
PRODUCT_GROUP_S_NAME	CIG				
CORP_UOM_CODE	STICK				
APO_UOM_CODE	Piece				
BRAND_HOUSE_CODE	1273				
BRAND_HOUSE_NAME	Smart		X		X
BRAND_HOUSE_S_NAME	Smart		X		X
LSKU_LOCAL_CODE	11553075-122	X	X		
LOCAL_SKU_NAME	Smart Blue 27	X	X		
PRIMARY_LOCAL_SKU_CODE	Y				
LOCAL_SKU_STATUS_CODE	1				
LOCAL_SKU_STATUS_NAME	Active				
LSKU_DATA_MAINT_GRP	119				
MDM Instance Name	Imperial Tobacco Finland		X	X	X

APPENDIX 8: PRODUCT DATA CONTENT IN MDM KALIDO (3/3)

Data field	Example value	NAV	Product Register	Siebel	Sinfos
RETAIL_SKU_CODE	2102832	X		X	
RETAIL_SKU_SHORT_NAME	Smart Blue LIP KS Box 27 (6.30 ND HW) FI				
PROD_SITE_ID	DE_LGH				
RETAIL_SKU_STAT_CODE	1				
STANDARD_SKU_SHORT_NAME	Smart Blue Smooth LIP KS Box 27 FI				
BRAND_PACK_VAR_SHORT_NAME	Smart Blue Smooth KS Box 27 Stick				
BRAND_VAR_SHORT_NAME	Smart Blue Smooth KS				
BRAND_SUB_FAM_SHORT_NAME	Smart CIG				
BRAND_FAM_SHORT_NAME	Smart CIG				
BRAND_HOUSE_SHORT_NAME	Smart		X		
PROD_GRP_CODE	01				
PROD_GRP_SHORT_NAME	CIG				
RET_PACK_TYPE_NAME	Box				
RET_BNDL_SHPE_TYPE_NAME	Stick				
RET_PACK_CONT_TYPE_NAME	27 pcs				X
PROD_TYPE_CODE	1				
PROD_TYPE_NAME	ITG Brand				
BE_NAME	Non ITG Cigarettes				
REP_BRAND_FAM_SHORT_NAME	Non ITG Cigarettes				
MANUFACTURING_COMP_NAME	Imperial Tobacco Limited				
TRADEMARK_OWNER_NAME	Imperial Tobacco Limited				
STRATEGIC_GRP_CODE	03				
STRATEGIC_GRP_NAME	Non ITG Brands				
MKT_CONSTRAINT_CODE	043				
MKT_CONSTRAINT_Name	Finland				
Reporting product group	Cigarettes			X	X
Technical product group	Cigarettes			X	X

APPENDIX 9: PRODUCT DATA CONTENT IN SINFOS (1/4)

	Sinfos	Where the data is found					
	Data field	NAV	Product Register	Siebel	Kalido	SQL	Notes
Identification/primary information	GTIN		X		X		for each hierarchy
	GTIN of base unit		X				
	Start validity date		X				
	GLN of data supplier					X	
	Name of data supplier		X				
	Target market					X	
	Sector					X	
	FOOD / NON-FOOD relevant part of profile					X	
	GLOBAL trade item					X	
	Action request		X				
	Reference level					X	
	Trade item unit descriptor						manually
	Non-public					X	
	Publication date		X				
	Start availability date time						manually
	Internal item ID of supplier	X	X				
	Number of base unit contained		X			X	combination
	Item name (long) (FI)		X				
	Item name (long) (SV)		X				
	Item name (long) (EN)		X				
	Item name (FI)		X				
	Item name (SV)		X				
	Item name (EN)		X				
	Despatch unit					X	
	Invoice unit		X				
	Ordering unit		X				
	Additional information					X	
	Bulk product					X	
	Consumer unit					X	
Identification/primary information	GTIN of item substituted for		X				
	GTIN of next lower packaking item		X			X	combination
	Number of next lower item contained		X			X	combination
	GLN of manufacturer					X	
	Manufacturers name		X	X	X	X	
	Content		X				
	Content: unit of measure		X				

APPENDIX 9: PRODUCT DATA CONTENT IN SINFOS (2/4)

	Data field	NAV	Product Register	Siebel	Kalido	SQL	Notes
Identification / primary information	Gross weight		X				
	Gross weigth: unit of measure					X	
	Net weight		X				
	Net weight: unit of measure					X	
	Base price declaration relevant					X	
	Base price relevant content	X	X		X		
	Base price relevant content: unit of measure				X		
	VAT-% (FI)					X	
	Brand name (FI)		X		X		
	Brand name (SV)		X		X		
	Brand name (EN)		X		X		
	Functional name (FI)		X				
	Functional name (SV)		X				
	Functional name (EN)		X		X		
	Product classification: Finland		X				
	GPC: Brick code						manually
	CPV classification		X				
	Pricing on the product		X				
	Freshness date on product					X	
	Dangerous goods		X				
	Dangerous substances		X				
	Material safety data sheet					X	
	Biocide					X	
	Variable measure trade item					X	
	Country of origin		X				
	Customs tariff number		X				
	Levying kind					X	
	Levying class		X				
	Import classification type					X	
Additional indication	Country of last provenance		X				
	Organic origin					X	
Product class specific information	Tobacco products: nicotine		X				
	Tobacco products: tar		X				
	Tobacco products: carbon monoxide		X				

APPENDIX 9: PRODUCT DATA CONTENT IN SINFOS (3/4)

	Data field	NAV	Product Register	Siebel	Kalido	SQL	Notes
Dangerous goos/Dangerous substances/Physical&chemical characteristics	Limited quantities					X	
	Class					X	
	Classification code					X	
	Packing group					X	
	UN number					X	
	Transport category					X	
	Technical name in transport papers: Language						manually
	Technical name in transport papers						manually
	Hazard identification number					X	
	Additional information for transport papers					X	manually
	Hazard symbols					X	
	R-phrase					X	
	S-phrase		X				hidden
	S-phrase		X				hidden
	S-phrase		X				hidden
	S-phrase		X				hidden
Logistics/Packaking/Price/Order	Packaking type		X				
	Packaking type code						manually
	Measurements: height		X				
	Measurements: height unit of measure					X	
	Measurements: length/depth		X				
	Measurements: length/depth unit of measure					X	
	Measurements: width		X				
	Measurements: width unit of measure					X	
	Bar coded					X	
	Is packaking marked returnable					X	
	Pallet loading height		X			X	
	Pallet gross weight		X				
	Stacking factor max		X				
	Storage stacking factor		X				
	Pallet type / code		X				
	GTIN-coded pallet					X	
	Pallet handling (one-way/re-usable)					X	

APPENDIX 9: PRODUCT DATA CONTENT IN SINFOS (4/4)

	Data field	NAV	Product Register	Siebel	Kalido	SQL	Notes
Logistics/Packaking/Price/Order	Number of despatch unit per pallet		X				
	Number of layers per pallet		X				
	Recommended sales price		X	?	?		
	Recommended sales price: currency		X				
	Handling instruction					X	
	Code		X				In PR, but manually
	Number						Manually
	Distance horizontal		X				In PR, but manually
	Distance horizontal: unit of measure						Manually
	Distance vertical		X				In PR, but manually
	Distance vertical: unit of measure						Manually
	Temperature (min): storage						Manually
	Temperature (max):						Manually
	Temperature: unit of measure						Manually

APPENDIX 10: SQL-QUERY FOR TXT-INPUT FILE FORMATION

```
SELECT
        dbo.TREK_Hinnasto.KulpakkEAN AS KEY_ALO_EAN,
        dbo.SINFOS_SIIRTO_STATUS.VoimaantuloPvm AS DAT_ALO_GUELTIGAB,
        '6410109999992' AS KEY_ALO_TEIL_NEHMER,
        'FI' AS KEY_PROFIL_LAND, NULL AS NUM_LFDNR,
        dbo.TREK_Hinnasto.Muutospvm AS DAT_GESPEICHERT,
        1 AS SWI_IMPORTIERT, 'FNF' AS KEY_PROFIL_SECTOR,
        0 AS SWI_TEST, 0 AS SWI_GEPRUEFT, 0 AS SWI_GESENDET,
        0 AS SWI_WEB_SENDEN, 0 AS SWI_KOPIERT,
        'ADMIN' AS KEY_LOGIN, NULL AS DATE_GESENDET,
        NULL AS DAT_EMPFANGEN,
        NULL AS NUM_ABI_VERSION, NULL AS NotInUse

FROM
        dbo.SINFOS_SIIRTO_STATUS INNER JOIN dbo.TREK_Hinnasto ON
        dbo.SINFOS_SIIRTO_STATUS.KulpakkEan =
        dbo.TREK_Hinnasto.KulpakkEAN AND
        dbo.SINFOS_SIIRTO_STATUS.MuutettavaEnari =
        dbo.TREK_Hinnasto.KulpakkEAN AND
        dbo.SINFOS_SIIRTO_STATUS.HinnastoId = dbo.TREK_Hinnasto.id
        INNER JOIN dbo.TREK_TuotePerustiedot ON
        dbo.TREK_Hinnasto.Prodsern = dbo.TREK_TuotePerustiedot.Prodsern

WHERE
        (dbo.SINFOS_SIIRTO_STATUS.Status = N'-1')
```

APPENDIX 11: NEW MR 6.0 REQUIREMENTS

Mandatory fields				
Attribute	**Description**	**BI**	**PI**	**PI**
Trade item unit descriptor	Describes the hierarchical level of the trade item. Trade Item Unit Descriptor is linkage between different levels of item hierarchy.	X	X	X
Start availability date time	The date when the trade item becomes available by a supplier.	X	X	X
Brand name (FI, SV ja EN)	Brand name can be international (eg. Coca-Cola) or defined individually for each country/market.	X		
Functional name (FI, SV ja EN)	Describes use of the product or service by the con-sumer. Functional name is one of the modules of a complete item description.	X		
GPC: Brick code	The GS1 GPC is a system that gives buyers and sellers a common language for grouping products in the same way, everywhere in the world in addition to their own internal and external classifications	X		
Is packaging marked	Specifies if the package is returnable.	X	X	X
Conditionally mandatory fields				
Optional fields				
Import classification type	WS2 data pool requires CN-Code and its Classification Type which specifies CN-Code type.	X		

APPENDIX 12: GPC-CODING AND STRUCTURE